VISUAL RECOGNITION OF WORDS

D.G. Bouwhuis

PROMOTOREN: PROF. DR. W.J.M. LEVELT
PROF. DR. H. BOUMA.

VISUAL RECOGNITION OF WORDS

Proefschrift

ter verkrijging van de graad van doctor in de Sociale Wetenschappen aan de Katholieke Universiteit te Nijmegen op gezag van de rector magnificus Prof dr P G A B Wijdeveld volgens besluit van het college van dekanen in het openbaar te verdedigen op woensdag 6 juni 1979 des middags te 4 00 uur

door

Dominicus Gabriël Bouwhuis

geboren te Leeuwarderadeel

Greve Offset B V - Eindhoven

Acknowledgements

I am very much indebted to the Institute for Perception Research, Eindhoven, for providing facilities to carry out the experimental and theoretical work reported here. The support of Philips Research Laboratories, Eindhoven is also gratefully acknowledged.
This research formed part of a general research programme into reading initiated by Dr. H. Bouma.
Many thanks are due to A.L.M. van Rens and in particular to J. Jacobs for their conscientious experimental work and for their assistance in data analysis.

Chapter 2 has been reprinted by permission of the copyright owner, the International Society for the Study of Attention and Performance. Copyright of Chapter 3 belongs to the Plenum Publishing Corporation. Parts of Chapter 6 have been reprinted from Perception and Psychophysics, the copyright belongs to the Psychonomic Society, Inc.

G.E. Luton of the Philips Translation Dept. and S. Marcus of the Institute for Perception Research contributed significantly to the final editing of the English text.

CONTENTS

Page

1 THE PROCESS OF WORD RECOGNITION AND READING 3

1 Reading in society 3
Disability to read 7
The case for research in reading 11
2 Scope of the present study 11
The reading process 13
Word recognition and word knowledge 15
The choice of experimental conditions 16
3 The global and analytic view of word recognition 17
Cattell's whole-word theory 17
Erdmann and Dodge's general word shape 20
Dominant letter theories 25
Decline 27
Application and education 28
New Look 28
Word length 29
Word frequency 29
Iconic memory 30
Pattern-unit theories and the word superiority effect 31
Reicher's theory of meaningfulness and word recognition 31
Johnson's pattern-unit theory 34
Redundancy theories of word recognition 36
Massaro's redundancy theory 36
The redundancy theory of Estes 39
The redundancy theory of McClelland and Johnston 43
References 47

2 A MODEL FOR THE VISUAL RECOGNITION OF WORDS OF THREE LETTERS 53

Introduction 53
Reading versus recognition 54
Word contour 54
Luminance 56
Duration 56
Type font 56

	Page
Eye position and visual interference	56
The recognition model	58
An experimental test	60
Results	61
Correct responses	61
Letter recognition and completion	62
Discussion	63
Word frequency	64
Interaction and independence	64
Extension of the model	65
References	66
3 WORD KNOWLEDGE AND LETTER RECOGNITION AS DETERMINANTS OF WORD RECOGNITION	67
Word recognition	68
Word counts	68
Shaping a vocabulary	70
Recognition and the role of knowledge	71
Experiment	73
Material	73
Presentation	74
Subjects	74
Results	74
Response times	74
Accuracy measurements	78
Discussion	81
Orthographic structure	82
Word frequency	82
References	84
4 THE CONTRIBUTION OF LETTER STRUCTURE TO THE FAMILIARITY OF DUTCH THREE-LETTER WORDS	87
Letter and digram frequency	89
The relation between letter frequency and decision latency	92
The relation between letter frequency and decision accuracy	97
The effects of digram frequency	98

contents continued

	Page
Discussion	99
A processing mechanism for lexical access	101
Phonological recoding	102
The contribution of single letters and digrams	103
Reference note	107
References	107
5 A RETEST OF THE WORD RECOGNITION MODEL WITH A VARYING KNOWLEDGE VOCABULARY	111
The relation between lexical decision and word recognition	114
A recognition model with a varying knowledge vocabulary	115
Results of the retest	117
Correct word scores	117
Letter scores	118
Discussion	119
Lexical decision versus word recognition	120
Subject differences in word knowledge	120
Subject differences in recognition	121
Conclusion	121
References	123
6 A REVIEW OF RELATED THEORIES AND EXPERIMENTAL DATA	125
Alternative word recognition models	127
The logogen system	127
The multicomponent model	128
Stages of processing	130
A Review of related experimental data	134
Visual interference and letter recognition	134
Retinal locations	137
Completion and redundancy	138
Frequency effects	140
Word knowledge	141
Reading research and education	144
Readability research	145
The temporal dimension in reading	145
The temporal dimension in reading education	146

contents continued

	Page
Dyslexia	147
References	148
APPENDIX: A MATHEMATICAL EXPOSITION OF THE LOGOGEN MODEL	155
Reference	158
SUMMARY	160
SAMENVATTING	163
Curriculum vitae	166

VISUAL RECOGNITION OF WORDS

1 The Process of Word Recognition and Reading

There is nothing strange about the fact that one can read these words. However, in this introductory chapter an attempt will be made to show that it is in fact most remarkable and, further, that in several respects, it may be considered as a privilege. These statements are illustrated by a summary treatment of the reading process in which a review is given of its various requirements. Especially the role of word recognition in reading is discussed. Basic components of the word recognition process will be employed in the following chapters to develop a model for the recognition of three-letter words. In this model it is made explicit how sensory information is extracted from the words of a text and how this information is complemented by word knowledge.

In the third part of this introduction we explore the evidence for the basic assumptions of the model, concerning perceived letter information and word knowledge.
Theories of word perception are as old as psychology as an empirical science, at the beginning of which reading was actively explored. It will be seen that, continuing into present-day theories, a distinction can be drawn between analytic, letter based, and global, whole unit theories of word perception. A general conclusion is that evidence in support of letter based word perception is quite strong. Nevertheless an essential property of constituent letters seems to be their positional information, which reflects global aspects of words.

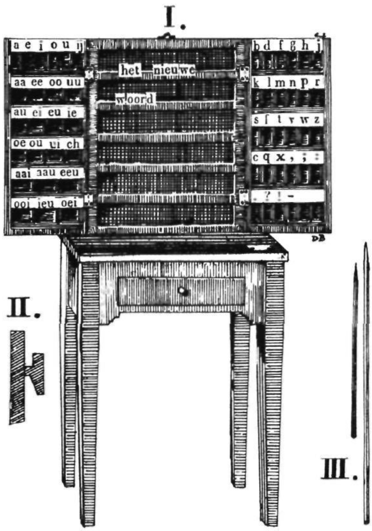
1 READING IN SOCIETY

The average reader is usually not aware of the complexity of the processes involved in reading. Reading strikes one as a routine activity, an activity as normal as eating or walking and so easily performed that it seems fruitless to ponder on its behavioral basis. Still, any experimental psychologist, by some strange coincidence not familiar with reading, and faced with its many requirements, ranging from the recognition of variously shaped ink impressions to the comprehension of implied meaning would state with unusual confidence that reading is just not possible.

It is a sad fact that until about 1800 most school teachers would have shared this view, at least for the greater part of their unruly pupils. Learning to read started with mastering the names of the individual letters in several widely differing type faces, and the reading of words started only after a few years of experience. Frequently, reading had not been mastered at all at the time the children were taken from school to contribute to the family income (de Booy, 1977). In the Netherlands it was only after the Education Act of 1801, which abolished doctrinal education, that new teaching methods for reading were introduced. Specially designed equipment, like the 'reading machine' of Prinssen, (Fig. 1) helped to reduce the time needed for acquiring a basic reading skill from several years to mere weeks. In other European countries improvement in reading education took place at about the same time.

Figure 1

The 'reading machine' developed by the Dutch teacher P.J. Prinssen shortly after 1800. It was used widely in the new reading methods and could be found in almost any self-respecting school in the Netherlands. Its purpose was to speed up the reading of words instead of concentrating on letter names. With the letters and letter combinations all words could be easily assembled in print-like quality at a time when the blackboard was hardly in use. A cross-section of a letter (II) shows its attachment.
Note the absence of the y, which is rare in Dutch, and the two characters for the letter s.



The common theme is the reduction of the number of different type faces, the early reading of words as opposed to spelling out single letters (Fig. 2) and the emphasis on comprehension of text. From that time on much more attention and effort were spent on the design of reading methods and illustrative material (Figures 3-6). Slowly changing views also led to different ways of teaching: analytic, letter-based methods of reading and global, word-based methods followed each other periodically (Chall, 1967). Nowadays a combination of both methods is popular, it is known as the structure method.

Even though basic reading skills can be acquired quickly, practice in reading still remains important throughout the years of schooling. In higher education the reading of foreign languages is required. Most scientific and literary material, constituting the main body of knowledge and culture at our disposal is in the form of reading matter.

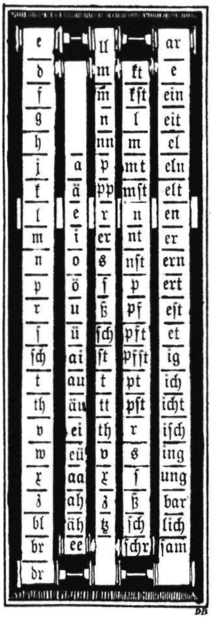


Figure 2

A German reading machine with letters in the Gothic type face. Here the letters were printed on moving belts which could be arranged to form arbitrary words which were visible through an opening in the front cover (not shown). In the picture the word 'lernen' (to learn) is displayed. Some letters were provided with marks, (m̃ in the middle column) next to the vowel mutating Umlauts. They indicated special articulation, for example dorsal or apical with respect to tongue position. (Picture adapted from Huey, 1968)

This fact and the long period of education probably account for the high esteem which reading enjoys. Though concerned with the same language, reading is much valued above speech; listening is rated below reading in many respects. Reading is not subject to criticism as is watching television, looking at comic strips or talking on the telephone, although the linguistic context may be identical. Furthermore reading is usually at once associated with books, less often with newspapers or written notes, and even less with subtitles on the television screen. Yet to take one example, elaborated nicely by Perry and Aldridge (1967),

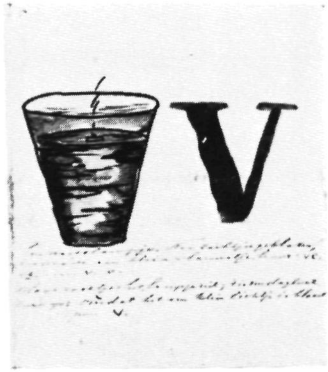


Figure 3.

A candlestick with candle burning brightly. If you blow, the flame drifts aside. Blow strongly, fff...

Figure 4.

A nightlight. Now blow softly, it's only a tiny little flame, listen ve, ve ...v

Figures 3-6 show examples of a letter reading board which was popular in the Netherlands round the turn of the century. The letter-sound correspondence was exemplified in a number of pictures, with instructions for the teacher on how to teach the letter names. The examples were sometimes rather artificial as demonstrated in the pictures for v (Fig. 4) and aa (Fig. 6). The pupils were encouraged to imitate the sounds and the appropriate gestures.
(Photographs by courtesy of the Historic Regional Museum Kempenland, Eindhoven)

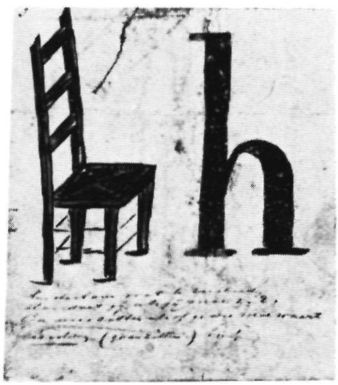


Figure 5

A chair to rest a bit. How are you doing when you're tired? Sit down as if you were very tired, ha ..h,h

Figure 6

Two girls are looking for flowers in the grass for mother's birthday. Finally they find a flower and laughingly call aa!

graphic illustration has a much longer history than written words. Prehistoric man could draw several millennia before he could write. A child can recognize a picture long before he can read words. But words can obviously communicate other ideas than pictures and do so in a much different way. For a long time it was the only way to record the spoken word. Written words can convey ideas of a highly abstract level, and they can summarize concepts and meanings while forming their own registered encoding, visible to anyone. Written and printed matter is probably the most efficient and cheapest memory around.
In addition, many more people are good at writing than at drawing, which makes text an easier means of communication than pictures. But it must also be acknowledged that the use of text has its limits, notably in very concrete information. Maps and blueprints can hardly be imagined as verbal descriptions. The increasing popularity of television, film and photography, the existence of radio or audio communication in general, certainly show an appreciation of non-textual media and spoken language. But whether it is overrated or underrated, it cannot be denied that reading fulfils a most central role in present-day communication.

DISABILITY TO READ

This section explores possible causes of the inability to read. Education, physical condition, specific reading disability and the writing system will be treated in succession and some of the consequences outlined.

In strong contrast with the subjectively unproblematic nature of reading is the observation that many people cannot read. In the fifties half of the adult word population was wholly illiterate. If normal reading ability is expressed as being at least at 4th grade level, then two thirds of the adult word population was wholly or semi-illiterate (Eisenberg, 1966). Lack of education is obviously the main determining factor which illustrates dramatically that teaching is indispensable for learning to read. Nevertheless, Western Europe counted at the same time 8% illiterates (Vernon, 1971) and the United States 11% (Eisenberg, 1966), mainly among the poor and black population. In the last decades increased school attendance and improved teaching methods in a number of countries have decreased the illiteracy rate to such a small proportion, that official figures are sometimes no longer given. The last figure in the Netherlands, based on male army recruits dates even from 1938 and amounts to .07%. There may also be political and social reasons for the publication of somewhat deflated figures, giving the impression of high educational success. So, the recent discovery that between 1% and 4% of the adult population in the Netherlands is analphabetic (Hammink and Kohlen, 1977) was met with disbelief and shame, despite the vagueness concerning the incidence. Hammink and Kohlen (1977) show that their estimates cover recent estimates obtained in England, Sweden, Belgium, France, Poland and the USA. The figures can be qualified by noting that there is a small group (.5% - 1%) which is basically illiterate, unable to read a word, while the remaining group may be considered semi-illiterate, unable to employ reading as a useful tool (Longley, 1976). It may be noted here that illiterates can sometimes recognize common verbal units like sign-boards, brand names, newspaper names or can infer them from situational context. Numbers seem in general to present little difficulty (Poelwijk and Verdiesen, 1978). From the literature sources mentioned it appears that irregular school attendance, frequent illness in youth and adaptation problems are typical for the majority of the cases studied. Only a small number of them ever participate in reading courses; many are very reluctant to admit illiteracy and resort to clever techniques in concealing it. Frequently learning progress is slow and the resulting disappointment may be a reason to give up prematurely which, in turn, leads to increased frustration. Nevertheless, many illiterates are observed to have developed a very good memory (Bolhuis, 1978), probably more or less forced by their inability to consult written material.

Reading disability may also be constitutionally determined, e.g. by sensory defects, like low vision or blindness; intellectual defects or brain injury, which may lead to acquired dyslexia. There are reliable statistics concerning the incidence of acquired dyslexia[1]. These are lacking for reading disability due to congenital defects.
These cases are usually not contained in statistics on education because normal school attendance is often impossible. A number of studies, cited by Andrew (1978) indicate that reading disability is dramatically prevalent in delinquents. The disability is strongly related with poor verbal skills. Nevertheless, reading the individual case histories of severely backward readers, it is not surprising that maladjustment may result. Finally, specific reading problems are frequently encountered among children. This so called specific reading disability is referred to variously as (congenital) word blindness, legasthenia, reading retardation but most often (developmental) dyslexia (Eisenberg, 1966). Phenomena connected with dyslexia seem to be confined largely to reading but a single cause is not known. There are indications of a deficient visual code to sound recoding and vice versa (Bouma and Legein, in preparation). Otherwise dyslectic children seem normal in every respect, also in visual acuity. Figures on the occurrence of dyslexia are rather variable and often unknown; sometimes they include also backward readers. Tarnopol and Tarnopol (1976) found from an international survey that percentages ranged from 1% in China to 33% in Venezuela, with a median value of 8%, all figures pertaining to children attending school. Specific estimates for the USA are 10% (Thompson, 1966), for England 5.7% to 15% (Vernon, 1971) and for the Netherlands 5% (Bouma, Legein & van Rens, 1974). The common problem of alphabetic languages is that children are forced to divide up their language units in phoneme-like segments. This is not accomplished in an easy or automatic fashion. It has been suggested that dyslexia occurs less often in languages where there is little discrepancy between phoneme-grapheme

[1] the brain injury literature also suggests that numerals and letters are processed differently.

relations than in languages with an irregular pronounciation, like English or French (Money, 1967).
Recent results show this not to be true; figures for Spanish, a language with high phonetic regularity, are lacking but reading problems in Latin America seem more frequent than anywhere else (Critchley, 1970; Gibson & Levin, 1975; Tarnopol & Tarnopol, 1976). Dyslexia has been shown to exist in almost any language; Italian, Finnish, Hebrew, German, Russian and Arabic to name only a few. Gibson and Levin (1975) and Tarnopol and Tarnopol (1976) concluded that teaching quality is much more directly related to the occurrence and nature of reading problems than type of language. There is, however, another claim that dyslexia is very rare in languages with an ideographic script. Except for case histories supporting this claim not many figures have been published on the issue. For Chinese the official dyslexia percentage was already mentioned to be 1% which is indeed very low. In a report by Butler (1976) it is noted that dyslexia as it occurs in the USA does not exist for Chinese. There is a much cited study by Makita (Gibson and Levin, 1975) from which it appears that dyslexia in Japanese 'kanji', an ideographic script, has an occurrence of less than 1%, as opposed to 'kana', a syllabic script in which dyslexia is much more frequent. Some other indications are given by Steinberg and Yamada (1978).
According to the survey of Tarnopol and Tarnopol (1976) dyslectic children respond well to remedial teaching which stresses the developmental character. Nevertheless, it is surprising that dyslexia is very rarely associated with adult illiteracy. Neither the comprehensive review of Hammink and Kohlen (1977) nor the specific study of Poelwijk and Verdiesen (1978) even mention the phenomenon. Yet, many of the reading errors of semi-illiterates, like vowel shifts, lengthening of words and severe problems on long words, resemble those of dyslectic children.
It is clear that without remedial teaching the dyslectic child is doomed to semi-literacy in adulthood.

THE CASE FOR RESEARCH IN READING

The preceding section has attacked the notion that reading is easy and within the possibilities of everyone. For successful reading and comprehension of text, extensive training, suitable reading material and good quality education is required. But at present even the best educational reading programmes for alphabetic languages cannot decrease the rate of dyslexia to substantially below 5%; the lowest reliable figure for Western European countries given in Tarnopol and Tarnopol (1976). It is, therefore, possible that a value in the neighbourhood of 5% represents the minimum attainable. The quality and effectivity of teaching methods can at present only be evaluated by the end stage of the reading process and comprehension, and hardly by insight in the functional processes of reading. Therefore, an educational success score in the vicinity of 95% which is representative of 'good' educational systems in a number of countries can by all standards only be regarded as impressive. But it should be realized that each failure represents considerable difficulty in the visual communication process which has been adopted for society as a whole. As long as reading is valued in society as it has been until now, the need for research on reading and its various aspects is all the more pressing.

2 SCOPE OF THE PRESENT STUDY

Reading involves a rather large variety of component processes. Many of these can separately, and justifiably, be studied and still be relevant to reading. But as a short review of the reading process will show, e.g. the perception of printed words involves a number of conditions which makes generalizability of separate studies to reading and vice versa difficult. This calls for the study of component processes under normal reading conditions, and this is the procedure followed in the present study. The specific object under consideration is the recognition of words during reading. After a review of the reading process, with special emphasis on word recognition, an introduction is given to the experiments reported here. The studies are specifically intended to clarify the role of letter recognition in the perception of words on the one hand, and on the other hand to obtain information on the role of the reader's word knowledge. Next the experimental procedures are outlined. The results obtained are explained in an explicit theoretical framework and as a consequence the assumptions on which the theory rests receive careful consideration. The assumptions are mainly concerned with an issue which has its analogue in reading education: the distinction between analytic and global perception.

The analytic theory holds that letters mediate word perception; the global theory assumes that the word is processed as a whole. The results of early and important studies in the psychology of reading, some of which go back to 1880, were taken to support the whole-unit theory of word perception. These studies carried out by Cattell, and Erdmann and Dodge (Woodworth, 1938) are examined in more detail. It is shown that the classical interpretations of the results are in all probability wrong and that global perception was induced by the procedures employed. Word perception theories stressing the importance of letters were quick to follow and these are briefly summarized. A period of relative neglect ensued in which a shift can be observed towards more fundamental and theoretical concepts and on the other hand towards applied reading research.

After the second World War, a number of new concepts revived interest in the study of word recognition, which eventually led to renewed interest in reading. The analysis of the role of letters and of word units in word perception was much refined - one might even be inclined to say too much refined. The main visual effects as observed on the role of letters and knowledge of words are therefore presented here in a detailed way. They permit the conclusion that many different experiments indeed substantiate the assumptions basic to the model of word recognition developed in this thesis.

Although the theoretical standpoints sometimes give the impression that there is a continuous struggle going on between letters and whole words, it will appear that usually the reader wins.

THE READING PROCESS

There is only a limited area in our field of view in which we can see sharply. This area corresponds to the fovea of the retina, and its center - where visual acuity is highest - corresponds to the direction of gaze, or the fixation point. Objects that do not appear in the foveal area are said to be eccentric; when they are just outside the fovea the term parafoveal is also used. Since reading requires the detection of relatively small features in a dense arrangement, the part of text to be read has to be brought into the foveal area. This necessitates movements of the eyes in such a way that the point of fixation reaches the part under attention. The eyes do not move in a continuous and smooth fashion over the lines but do so in jumps, called saccades. Eye saccades were first described by Javal (Woodworth, 1938) who watched the eyes of reading children. The extent of an eye saccade is variable but averages 2° visual angle, corresponding to about eight letter positions at reading distance. The duration of a saccadic movement averages 20 milliseconds (ms). When the eye movement halts, the ensuing fixation pause lasts on average 200 ms (Bouma & De Voogd, 1973) but may be as short as 50 ms. Fixations seem to be distributed in an almost random fashion over the lines of print; they may fall on any letter of the words or on spaces between them. Many words are thus seen in indirect vision at some distance from the fixation point.
Visual acuity is lower in those areas, but an other phenomenon is introduced as well. There is mutual interference between the adjacent letters of a word, which impedes their perception. This effect increases strongly with distance from the fixation point (Woodworth, 1938; Bouma, 1970). Thus, although some words can be clearly perceived during reading,

there are others which are not directly looked at and which are not brought into the foveal area by control of eye movements.

When the eye moves over the line a certain word enters the (useful) visual field from the right. After one eye saccade the fixation point has approached the word, or may even fall on it. In the next pause the word may have been shifted to the left of the fixation point and will move out of the visual field after more saccades. One word may thus be seen several times but usually in indirect vision. Yet all these presentations give the impression of just a single word that is presented only once. This need not be surprising in view of the fact that we are usually unaware of our eye movements during reading but also in other looking tasks. Apparently the information which may be extracted from the word upon each 'presentation' in the visual field reaches invariably the same central representation of the word and probably strengthens it continually.

As has been mentioned before, there are basically two theories of the process by which the word is recognized. According to the first the letters which are recognized evoke the representation of the word in memory. In this way the meaning which was associated with the same letter pattern on many former occasions will be retrieved and made available. The other theory holds that letter analysis is unnecessary and even undesirable. It is assumed that the word itself functions as a unit, just as a single letter or another visible object is a unit. The letters may mediate its global outline, its length and characteristic structure but lose their perceptual identity in the arrangement they form.

In both theories it is sometimes suggested that for the process of evocation, a phonological code is produced from the perceived pattern which, in turn, evokes the associated meaning.

Evidently, for reading aloud an articulatory coding of the word has to be produced at some point for the spoken utterance.

Every time the word is seen when the eye moves over the line it is preceded and followed by the same neighbouring words. As soon as recognition of these words sets in and the meaning comes through, a semantic and syntactic analysis has to ensue to render the words comprehensible in their context. Although important and interesting processes are involved in comprehension, including that of short phrases, they fall outside the scope of the present study and will not be pursued any further. Here it will only be examined to what extent the recognition of single words can be described with elementary visual factors and with knowledge of words.

WORD RECOGNITION AND WORD KNOWLEDGE

In the following chapters a quantitative model for the recognition of Dutch three-letter words is developed and tested. This model describes perception of a word as originating from recognition of the constituent letters. Letters may be perceived imperfectly owing to conditions inherent in reading. As a result errors may be made, leading to letter confusion. When the wrong letters belong to another word they will evoke that word, resulting in an erroneous response. This is why the model is called the letter confusion model of word recognition. Basically the model describes an analytic, letter-based operation. The emphasis on letter position within the word reflects the fact that global properties are also taken into account. Perception of letter strings, however, is only half the story. Since words are known they form part of the reader's memory. It is assumed that a perceived letter string, forming the letters of a word will retrieve that word and the corresponding meaning from memory. Consequently, whether a word will be perceived is critically dependent on its representation in memory. This model prompts two types of experiment, one in which recognition of words is studied as a function of the identifiability of the letters and one in which knowledge of words is investigated.
The first type of experiment studied the recognition of

letters in different positions of a three-letter string, and the recognition of words under the same conditions. Knowing how letters in various letter positions are perceived permits one to predict which letter strings may be perceived for a presented real word. Then, if word knowledge is available, the perception of specific words can also be predicted.

The second type of experiment, therefore, studied the reader's word knowledge. It employed a lexical decision task which involved all Dutch three-letter words. Words appeared to vary substantially in familiarity which was not only caused by their differential frequency of occurrence. Visual factors too, residing in the constituent letters, contribute to their familiarity. Since familiarity may be expected to affect recognizability of words, the model was also tested with a word vocabulary derived from the lexical decision study.

The choice of experimental conditions

In all experiments the stimulus, whether a letter string or a word, was presented only once. This corresponds to what is seen in a single fixation pause of the eye. What is seen in this single presentation is therefore not influenced by previous presentations at other places in the visual field, nor can report be affected by later presentations. This implies that performance would be inferior to that in normal reading, but there the word is surrounded by other words interfering with its perception. Yet the chosen procedure simulates a relatively limited part of the process of word recognition.However, if one wishes to investigate which visual properties of a word mediate its recognition other words must be excluded from the presentation.

The choice of other parameters of the presentation will be discussed in the experimental reports.

A final word concerns the nature of the responses given by the subjects in the experimental tasks. In word recognition the reader pronounces the word that he perceived, which is a natural situation. It is less natural to pronounce a letter from a string, though it may be argued that it occurs in abbreviations. In a lexical decision task, however, the subject says whether the letter string is a word or not, and such a task is rather uncommon. A more natural alternative method that would be equally effective for the purpose did not appear to exist. The generalizability of the results obtained with this method is also discussed in the report concerned (Chapter 3).

3 THE GLOBAL AND ANALYTIC VIEW OF WORD RECOGNITION

Research on reading was one of the first areas of experimental psychology and reviews of the extensive body of research can be found in Huey (1908, 1968), Woodworth (1938), Neisser (1967) and Henderson (1977). In connection with the letter approach embodied in the word recognition model which is proposed here it is sufficient to retrace the role of letter perception in the various theoretical interpretations. Thus how probable is it that letter perception per se could account for the results of early experiments? Alternatively, do theories of word recognition allow for letter perception? Table 1 shows in a highly schematic fashion the main theoretical conceptions of early reading investigators together with some relevant experimental details. Two distinct groups appear, one centering on whole word perception, the other on parts of words. Starting with whole-word theories we shall examine which results led to the shaping of the theories, and why.

CATTELL'S WHOLE-WORD THEORY

Cattell maintained, perhaps as an understandable reaction to tradition, that familiar words are not read by spelling out all letters, but as wholes. First he found that naming times for letters were as long as those for words. By itself this result does not imply that words are processed as wholes, since letter recognition could also proceed in a parallel fashion. But it rules out the then common view that word recognition is based on spelling them out.

Table 1

Review of early experiments on word perception derived from Woodworth (1938), Huey (1908) and Wundt (1914).

AUTHORS	BASIS OR RECOGNITION	TYPE OF EXPERIMENT OR RESULT
Cattell, 1885	Word is read as a whole.	a) Reaction and naming times for letters and words. b) Span of apprehension for letters and words.
Goldscheider and Müller, 1893	a) The sound of the determining letters evokes the sound of the word. b) Sometimes the determining letters are visually completed.	Presentation of incomplete words; introspection.
Pillsbury, 1897	a) Form of word, word length and location of clearly seen letters.	Presentation of mutilated words.
Erdmann and Dodge, 1898	General word shape.	Perceptual conditions rendering letters indistinguishable. Reaction times.
Zeitler and Wundt, 1900	Dominating complexes, which are assimilated to words. Global features are important	Short presentation times, attentional instruction. Introspection, type of errors.
Messmer, 1903	Dominating letters, ascenders, upper half of lines, internal character of words	Distinction between 'letter' and 'word' readers
Huey, 1908	a) For most readers the total form, the length and the inner structure b) Total outline.	a) Type of errors. b) Introspection.
Kutzner, 1916	Length and number of projecting letters.	Eccentric presentation.
Grossart, 1921	Word shape in eccentric vision.	Eccentric presentation.
Schumann and Wagner, 1920	Several letters from all parts of the word.	Series of 15 unconnected letters and attentional instruction.

However, it is hardly ever mentioned that Cattell (Woodworth, 1938) adhered strictly to the subtractive reaction time scheme evolved by Donders (1869, 1969). According to this model, if letters were processed as units, the time for processing them should be additive. What Cattell (Erdmann & Dodge, 1898) did was to establish a 'central' time for processing a single letter and then to compare this time with that for processing a whole word, which turned out to be the same. Huey (1968) claims even that this time was appreciably less for words, which is indeed found a few times in Cattell's data (Erdmann & Dodge, 1898); on average, however, times for words are a few milliseconds longer than those for letters in comparable situations. Considering the reliabilities of the data the differences are negligible, especially since they are prone to measurement errors. Cattell measured the naming times with the Hipp chronoscope (Fig. 7) which was controlled by a voice switch of his own design (Fig. 8). Careful checks made by Ach (1905) showed that the uncertainty of this device was minimally 3 ms. when it was optimally adjusted and fitted with special control switches; it could have been much larger. Since this is of the same order as the time differences found, no great weight can be attached to them. An alternative explanation for time equality is that the letter units are processed in parallel;

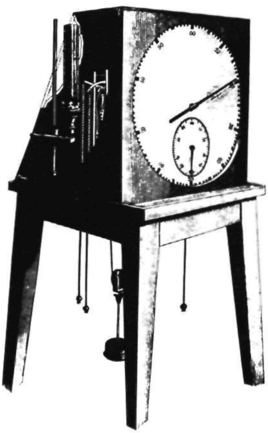


Figure 7 (left)

The Hipp chronoscope, used widely in reaction time studies. It permitted readings in units of 1 ms. Starting and stopping was controlled by electro-mechanical switches.

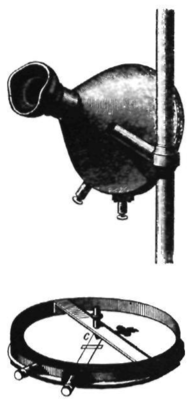


Figure 8 (right)

Voice key made by Cattell. The subject spoke into the mouthpiece (top) and vibrations of the membrane (bottom) closed the contact C, which triggered the stopping of the chronoscope.

this should lead to about the same processing times as those for a single letter. Time equality for words and letters is not a sufficient argument for whole word processing.
It may be noted here that Cattell indeed rejected some assumptions of the subtraction method on the basis of his results (Henderson, 1977) but maintained the whole-word explanation. A second argument Cattell gave for whole word processing (Woodworth, 1938) was that more letters could be reported when they were presented in a familiar arrangement (a long word) than in random strings. The line of reasoning should then be somewhat like the following. Subjects could report only 4.5 letters on average from an unconnected series, which was interpreted as the maximum number that could be perceived. Long words, of about 15 letters, could be easily reported in the same conditions, whereas 10 letters of all 15 could not have been perceived at all. Consequently, the word could not have been perceived as a set of letters, but only as a whole. It is generally agreed nowadays that this reasoning is wrong, though rarely in connection with Cattell's findings. Memory factors are held to be responsible for the limited number of reported letters when actually more have been seen as demonstrated by research of Averbach and Sperling (1961). There is sufficient reason to believe that in Cattell's experiments individual letters could indeed have been recognized.

ERDMANN AND DODGE'S GENERAL WORD SHAPE

The experimental results of Erdmann and Dodge (1898) supported the notion that not only was word shape the determining factor for word recognition, but even that letters did not need to be perceived at all. These findings were particularly hard to accept for their contemporaries (Huey, 1908, Wundt, 1911a). Wundt, after discarding the whole-word interpretation of his former assistant Cattell, stated on one occasion that Erdmann and Dodge's results were due to the long presentation time they employed (100 ms) and to the use of artificial light sources instead of daylight. On another occasion he assumed that they tried to approximate the reading situation too much. Huey (1908) argued that the majority of investigators disagreed with Erdmann and Dodge and pointed repeatedly to the carefulness of his own experiments.

Nowhere, however, was a real cause mentioned for the admittedly rather strong results of Erdmann and Dodge (1898), though the reason presents itself in the detailed and clear accounts of their own work, which was no less careful than that of several others. The probable reason is their tachistoscope. Following a specification given by Erdmann, Dodge designed what was to be the most flexible and efficacious presentation equipment for word stimuli (Fig. 9).

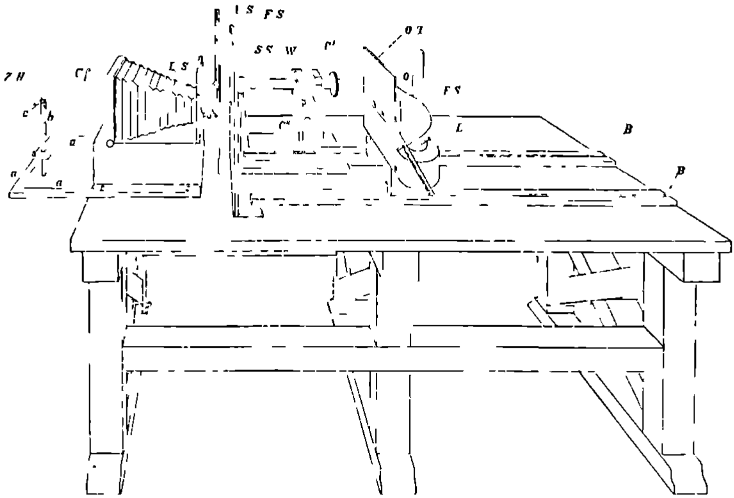


Figure 9.

Tachistoscopic equipment built by Dodge. The stimulus (O) is a transparency which is projected onto the screen G.f. The shutter S.S. is controlled by a chronoscope (not shown). Stimulus size was varied by sliding the stimulus over the rails B. Mechanical vibrations necessitated the heavy table, also designed by Dodge himself.

Technically it was indeed superior to the wheel tachistoscopes employed by Goldscheider and Müller, and to the fall tachistoscopes used by Wundt (Figures 10 & 11). The subjects could view binocularly, could accommodate far better and looked from a normal reading distance. Print size could be easily varied, presentation times were uniform over all parts of the display and could be accurately controlled. This was achieved by projecting the stimuli upon a viewing screen (Fig. 9), so the stimuli had to be constructed as transparencies. Erdmann and Dodge could obtain only a single Roman letter set, all others were of the Gothic type, which they thought to be not discriminable enough. This Roman type had to be reduced in order to approximate normal print size, but this could only be achieved by placing the stimuli with their projection light farther away from the viewing screen. This reduced the luminance, as did the ground glass on which the letters were pasted. They called in Dr. Dittenberger, who later participated as a subject, to measure the

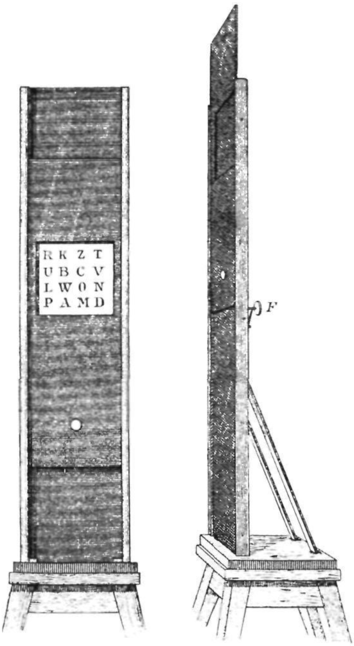


Figure 10

Fall tachistoscope for demonstration purposes as employed by Wundt (1903). Subjects fixated on the white spot (right) until the screen fell and momentarily exposed the letters.

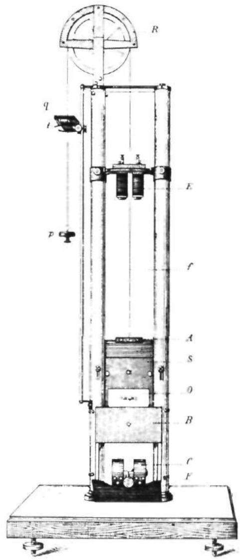


Figure 11

Fall tachistoscope for word recognition studies. The stimulus word 'Empfindung' is upside down, because the subjects observed through a non-magnifying astronomical telescope.

screen luminance, which turned out to be only .0325 of that produced by a Hefner lamp at the same distance as the stimuli. A Hefner lamp (Fig. 12) can be compared to a candle and its luminance output is by definition 1 HK (Hefner Kerze) which relates to other standards as follows:
1 HK = 0.9 international candle = 0.92 cd.

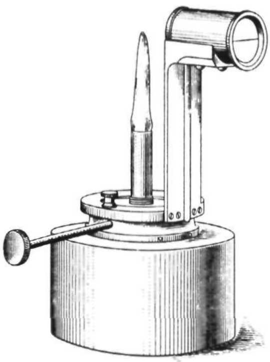


Figure 12

A Hefner lamp which was the luminance standard in Germany until 1948. The flame is 4 cm. high and burns on amyl acetate.

Thus it is not surprising that the subjects were dark adapted for 15 minutes and that they had to observe under a black cloth. The concomitant effect of such low intensities is a strongly reduced contrast sensitivity, making fine details imperceptible and impeding letter perception (Van Nes & Bouman, 1967; Patel, 1966). When the size of the letters was adjusted such that unconnected letters could no longer be perceived, words in the same conditions were perceived reasonably well. Consequently, on much better grounds than Cattell, they took the total visual form to be the basis of word perception. Unfortunately, the preparation of the stimulus words was rather cumbersome. The trouble of pasting single letters on thin glass sheets and aligning them carefully led Erdmann and Dodge to make only a limited number of them, initially no more than 69, later only 26. In their attempts to determine the viewing conditions, they used stimuli over and over again, so probably both author-subjects came to know which were the words they were preparing for each other. Disarmingly Erdmann notes that reading is much too routine an act to be trained by a few trials. For example, though Dr. Dittenberger participated in fewer trials, he made about the same errors as Erdmann himself. Now, it can indeed be induced from the data that

Dittenberger made the same errors as Erdmann, but he also made many more, which supports the effect of a fore-knowledge of possible stimuli and responses in the case of Erdmann. A final argument is provided neatly by the authors in an attempt to equate word perception as much as possible with single letter perception. To this end they learned at list of 26 words by heart, training intensively to realize the same associative structure as the alphabet. Even though, in the conditions applied, they could not perceive single letters, they correctly reported 94% of the words. In an earlier experiment with normally printed stimuli, presented at a large distance, continuously visible in daylight, they obtained a word score of 50%. Therefore, at the same size, but under much better lighting conditions, what were lesser known, printed words were far less well perceived. It is revealing that performance broke down on well-learnt short sentences, presented under the same conditions. After the identification of the most probable sources of Erdmann and Dodge's (1898) theoretical interpretations, the case of whole-word perception seems less strong than before. Making letters indistinguishable and increasing response redundancy necessitates word-feature or global-feature perception but does not imply that letters are denied any role in other conditions.

The criticism directed at Erdmann and Dodge (1898), however justified, was wrong and has probably led to a relative neglect of their otherwise clever and intriguing observations. In principle, limiting a vocabulary to 26 entries to equate them with single letters is a sensible control procedure, absent in many word superiority studies to be performed later. The authors discovered too, that more letters were perceived than were actually reported, for which phenomenon they presented two related explanations. First they postulated a visual (iconic) memory that faded during the report, and second, articulatory interference would tend to suppress responses still to be given. Both factors are standard interpretations today.

DOMINANT LETTER THEORIES

Essentially all other theorists mentioned in Table 1 adhere to the perception of dominant letters, or features directly produced by dominant letters, like word contour. Smaller features are not mentioned, except when referring to the inner structure of a word, which is strongly letter-bound. The most general theory of recognition is the association theory which was developed in the late 19th century by Wundt (1914), who borrowed several concepts from Herbart. The association theory influenced most investigators of the reading process quite strongly, notably Huey (1908). Therefore dominant letter theories will be discussed here in the framework of Wundt's theory, in connection with the findings of Zeitler who worked in his laboratory. The association theory embodies three main concepts: fusion, assimilation and complication. Fusion may occur in 'intensive' stimulation, like auditory stimuli, when tones fuse into a chord. It is significant that fusion was not thought to enter into word recognition, letters did not fuse into a word form, which would make them subordinate to global features. A complication arose when a visually perceived word evoked the auditory image of that word, its sound. The most important process, though, was thought to be assimilation and interaction between direct and reproductive elements. It was the connection between externally induced sensations and simultaneous memory elements (Wundt, 1911b). Wundt (1911a) discussed assimilation in particular in connection with the findings in Zeitler's word recognition studies. On the basis of the unanimity of the reports of the participating subjects Zeitler concluded that on the presentation of the word the dominant letters were seen in their structure. They were, then, immediately completed by the rest of the word. The interpretative difficulty for the present-day reader is that this act was perceived as simultaneous. It followed from introspection that dominant letters come first, and are simultaneously completed.

The idea (Vorstellung) and the sensation of the same object are not to be separated; recognition is composed of both aspects (Wundt, 1911b). It is therefore, difficult to say whether in early reading theory word recognition was thought to be mediated by letters or by more global features, like the structure of the dominant letters, which evokes the memory traces. Zeitler (Wundt, 1911b) held the view that whereas dominant letters seemed to determine where the indistinctive letters were placed in the word, the latter were not functional in word perception. They could even be 'suppressed' by the dominant letters, because the stimulus letters were sometimes shifted in position in the word responses and word length was not always preserved. Recently, Schiepers (1976) showed that word length can function in word perception, but also that length perception is not very accurate. On average a shortening of 5% may be observed. Woodworth (1938) remarks rightly that the contour of error responses obtained in Zeitler's experiments resembles closely that of the stimulus words. Finally, it is never clear whether a letter in a response word which is shifted relative to its position in the stimulus is also the one which has been perceived in that position. The chance level for guessing a vowel, for example, is 20%. So, when a guessed vowel appears in another position of the stimulus word, it cannot always be interpreted as a position shift; compare for example the response 'open' to 'upon'.
In 1923 Korte (Woodworth, 1938) details the effect which Zeitler called letter suppression. Korte (Woodworth, 1938) showed that adjacent letters interfered with each other in perception. This leads to the concept of lateral interference between letters, which would then seem to be a symmetric effect. Small letters interfere with dominant letters as well as dominant letters interfere with small ones, but dominant letters are more distinctive, because of size or features and can be reported more easily.

DECLINE

Surprisingly, Korte's finding is the point at which early reading research ends. It is the most recent reference in Woodworth (1938) which concerns fundamental processes in reading, a publication then 15 years old. In Germany, reading was not any longer considered to be a fundamental psychological issue. The handbook of Ebbinghaus and Durr (1913) treats the whole of reading research in a footnote. The writings of the 'Würzburger Schule', headed by Külpe, (Boring, 1950) hardly contain a single reference to word recognition. One even comes to consider it as ominous that Kulpe took the chair of Erdmann - who had published the volume on reading with Dodge (1898) - in Bonn in 1909.
Also in Gestalt psychology there was no place for word recognition; Gestalt psychologists emphasized the nativistic approach, while words, obviously, have to be learned.
In America the theorizing of reading research became suspect in behavioristic circles. It was common practice that introspective reports of subjects were accepted at face value, while statistical procedures were usually completely lacking. The number of subjects participating in the experiments was rather limited, and traditionally they were the authors themselves, or at any rate persons of a high educational level. They all had firm theoretical preconceptions, even though they were very careful observers. They considered words like 'Aufmerksamkeit' (apprehension) and 'Bewusstseinszustand' (state of consciousness) to be very frequent words (Wundt, 1911a).
In his history of experimental psychology, Boring (1950) hardly mentions Cattell's reading research and mentions Huey only as one of the 25 psychologists who took their doctorate with Stanley Hall. Dodge is just mentioned as the 27th eminent psychologist in the USA. Yet, besides his output as a reading investigator he invented the cornea reflection method of eye movement studies and the half-mirror tachistoscope, which are both still the most widely used types of equipment in their fields.

The second edition of 'Experimental Psychology' by Woodworth (Woodworth & Schlosberg, 1956) omitted the chapter on reading which appeared in the first edition, though bits of it were adopted in the chapter on attention. And the handbook of experimental psychology by Stevens (1951) does not contain a single reference to reading in its 1400+ pages.

APPLICATION AND EDUCATION

While the interest of psychologists shifted away from reading research, it was eagerly taken up by educationalists. Much reading research did not fail to be applied immediately in the field of reading education (Huey, 1908). This led to studies of visual fatigue and eye movements (Carmichael & Dearborn, 1947). Further, there was the search for the perceptual functions that make good readers, a riddle not even solved today. Finally, the readability of texts was studied, without much reference to the reading process itself.
So, maybe it is not surprising that applied reading research has not revealed much useful information on the reading process, though it may have helped to increase the awareness of causes for reading problems.

NEW LOOK

In 1948 Postman, Bruner and McGinnies found that personal values influenced the perception of words. Hunger would predispose someone to see steak for streak, or bread for broad. Solomon and Howes (1951) showed that these effects could be better or equally well described by the subjective probability of words. The view that perception is influenced by expectations became known as the New Look, and, since the views were not always identical, also as the New Looks. Good overviews have been provided by Erdelyi (1974) and Henderson (1977). Here we will mention two factors which are relevant in the present context. A review by Postman (1963) reports findings on the recognition of word as a function of the frequency of usage and the length of words, in visual and auditory presentations.

Word length

For visual presentations it is stated that longer words are more difficult to recognize, especially for decreasing word frequency. Postman (1963) argues that too little information can be extracted from longer words in a single glance. Such a statement is reminiscent of serial letter processing. Doggett and Richards (1975) have reanalyzed the data and show that the length effect presented by Postman (1963) is limited to the words of the lowest frequency only. Comparing the plots made by Doggett and Richards with figure 4a of Postman's review (1963) we can see that Postman's presentation is misleading in the sense that he implies that the length effect is present at higher frequencies as well. A further difficulty noted by Doggett and Richards (1975) is that in the few experiments obtaining a word length effect an extremely small number of words were employed. Replicating the experiments at three word frequency levels with words lengths ranging from three to eleven Doggett and Richards (1975) obtained no length effect at all. A moderate length effect was obtained for subjects of a lower verbal ability, but alone for rare words. Richards and Heller (1976) presented also pseudowords, next to real words, and found a very strong length effect on meaningless pseudowords. Richards and associates (1975, 1976) conclude that a length effect obtains only when the word is hardly known; in that case more information must be extracted from a long word than a short word to recognize it. The role of length in word perception will be treated more fully in the section on pattern-unit theories.

Word frequency

The word length effect, brings us to the other, much studied factor in the New Look: word frequency. Howes and Solomon (1951) found that the recognition threshold (number of presentations before correct response) increased strongly for low frequency words. Frequency effects have been found ever since, but there are two important cautionary remarks to make. First, the frequency effect is not as strong as it is often thought to be. This is pointed out clearly by Erdelyi (1974) for a number of studies. Second, as Richards (1973) has noted, in many cases where a frequency effect was reported, word frequency was actually 'constructed' by varying the number of presentations of (artificial) words. Richards (1973) showed that subjects tried to remember the presented words for the recognition trials later in the experiment and concluded that increased performance for the frequently presented words had no perceptual basis. What this all amounts to, is that, whereas a real word frequency effect is undeniable, the main effects of word recognition are visual in nature. But the New Look has certainly shown what the effects of word knowledge, or word availability, as reflected in word frequency, may be.

There is also an important methodological shift to be noted which was caused by improved psychometric methods. Perceptually relevant stimulus attributes were no longer inferred from the reports, but variables were chosen beforehand and manipulated systematically in order to study their effect on the recognition threshold. Furthermore, the threshold was a new concept in word recognition, as was the introduction of recognition probabilities as a dependent measure.

ICONIC MEMORY

In 1961 Averbach and Sperling showed convincingly that many more letters could be perceived at a single glance than 4.5 letters on average. Both articulatory interference and a fast decaying visual memory were held responsible for the limited number of reported letters. Averbach and Sperling (1961) arrived at an estimate of almost 10 letters, which even exceeds the first estimate of Erdmann and Dodge (1898) of about 7 or 8. The importance of this result for word recognition is that during and after presentation of a word, the search for that word can proceed from many more letter positions than 4.5. This can account at least partly for the superior performance of words relative to letters strings. Some criticism has been raised regarding the evidence for a visual short-term store with a large capacity (holding, 1975). Combining the available evidence, however, and considering the applied control procedures, one must conclude that the existence of iconic memory, as it has been called by Neisser (1967) seems to have been established fairly well (Coltheart, 1975).

PATTERN-UNIT THEORIES AND THE WORD SUPERIORITY EFFECT

The global view of word perception is stated most clearly in what might be called pattern unit theories. They assume either that simple visual features directly evoke whole word units without intermediate letter perception, or that meaningful words contain higher order features that meaningless letter strings do not have.

Reicher's theory of meaningfulness and word recognition

Pronouncing a word is only one way of unifying the letters in a composite pattern. Other perceptual ways might exist and the main problem is in what stage of the recognition process the integration actually occurs. It took almost 80 years before another whole-word theory was proposed after Cattell published his accounts on the subordinate role of letters.

Reicher (1969) showed in an experiment that letters were more perceptible in words than either in meaningless strings or when presented alone. Figure 13 shows the experimental conditions. The subject was asked to report the letter in the indicated position of the presented stimulus pattern. There were always two alternatives given, both of which formed words in the case of word stimuli. The subject could not, therefore, employ his word knowledge to choose the correct letter of the two. Since letters in words were perceived more accurately than in nonwords or more acurately than single letters, Reicher (1969) concluded that it was the word context that made the perception of its letters easier. Meaningfulness of words would then tend to increase the perceptibility of letters. Wheeler (1970) repeated the experiment with some changes in procedure. The stimulus conditions are shown in Fig. 14.

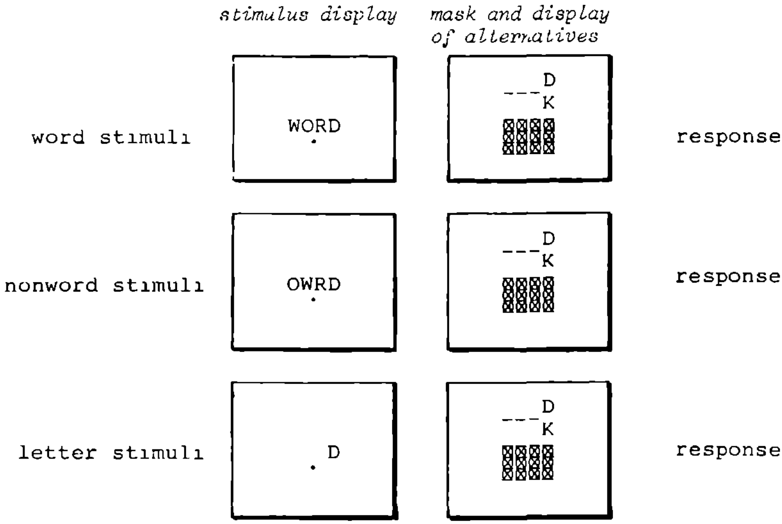


Figure 13

The stimuli in Reicher's experiment (1969). The point indicates the fixation point; in some conditions there was also a stimulus below the fixation point. Single letters could appear in each of the eight positions occupied by the words.

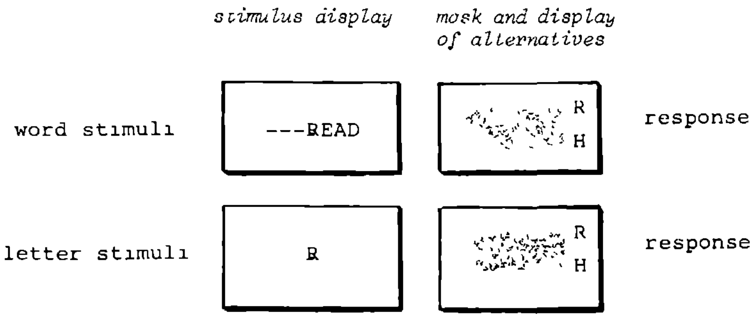


Figure 14

Stimuli in Wheeler's experiment (1970). The fixation point was directly below the fourth letter position. Words could shift in position as indicated by the dashes, which were not part of the stimulus.

In his experiment, too, perceptibility of letters in words was better than that of single letters; nonwords were not presented. Wheeler (1970) offered the following three explanations for the effect.

More features. It is possible that more features are extracted than only those pertinent to one letter. Henderson (1977) calls them transgraphemic features, those features which might be induced by the letter sequence. Wheeler (1970) mentions CO as a rounded group, and NI as a square group of letters. Space between letters might also be operative. Thus these features might be called global in the sense that they are not strictly related to single letters.

Feature selection. It is also possible that features are not passively extracted, but that after one feature has been detected a specific feature is selected from the stimulus pattern. The recognition system would select features in an interactive way. Essentially, such a system is even more global than the previous one. Selection of features would be guided by higher order units.

Coding with information loss. The third possibility Wheeler (1970) proposes is that the system tries to code the perceived pattern verbally, but that only one code is formed at the time of decision. The effect of this coding is that all other information is lost: when the coding is wrong an error will result. In the case of a word stimulus, however, imperfect perception might be compensated by the rest of the word, since there are more constraints for coding than in the case of a single letter. Henderson (1977) remarks that this procedure may be more general than just verbal coding and proposes memorability: when the subject succeeds in finding a code for the stimulus, which is easier for meaningful words than for single letters, he can memorize the perceived information better. This last argument differs from the other two in that feature perception is assumed to be equal in the case of words and single letters, but that redundant (or memorizable) information simply restricts the number of possible responses, producing higher scores for words.

Johnson's pattern-unit theory

One of the most detailed global feature theories was put forward by Johnson (1977). In this so-called pattern unit theory elementary visual features are directly mapped onto higher order units like words, which is compatible with the whole word theories of Cattell and Erdmann and Dodge (1898). The basic experiment that provides evidence in support of this system is according to Johnson (1977) the following.

Table 2

Stimulus conditions and response times in the study of Johnson (1977)

target	stimulus	response time	
		Yes-response	No-response
A	READ	539 ms	590 ms
READ	READ	514 ms	541 ms
A	A	493 ms	543 ms

A target is given before each trial and the subject has to report whether it is contained in the presented stimulus. Johnson considers only response times, since performance is almost errorless. It appeared that it takes subjects longer to report that an A is present in READ, than to see that READ is equal to READ or A to A. The explanation is that to detect the presence of an A in READ, the encoding of READ must first be parsed into letter components, because comparisons in working memory can only be performed on similar codings.

The parsing, thus, takes extra time, which is reflected in the longer response times. A first difficulty is that the observed time difference is rather small; it takes only 25 ms longer to see A in READ, than READ itself. Formally, the time needed for A-A and READ-READ should be the same, because they require one encoding and no parsing, still A-A is 21 ms faster. Another difficulty with the pattern-unit model is that in this experiment a word inferiority effect is predicted. The results indicate that identifying an A in READ is slower, and therefore less efficient, than identifying an A as A, which is contrary to the findings of Reicher (1969) and Wheeler (1970). Along the same lines it might be argued that when a nonword was presented the parsing into separate letters would be easier and faster than in words, and consequently identifying a target letter would tend to be faster in nonwords than in words - again a word inferiority effect.

A further type of experiment relevant to pattern-unit theory concerns the identification of words varying in length. Johnson (1977) mentions three experiments in which the subjects had to decide whether the stimulus word matched a predesigned target or not. The lengths employed were 3, 4, 6 and 8 letters and there were no significant differences between identification times for different lengths. The residual time differences still found between the averages for lengths are, however, between 20 and 30 ms. These may be sufficiently small to prevent them reaching significance. The problem is that the interpretation of the earlier results (Table 2) rests completely on differences of this size, relating to exactly the same comparison processes. The argument supporting the pattern unit theory is that no length effect exists, because the unitary encoding is independent of length, or the number of letters. The equality of response times can, however, also be explained by a theory holding that letters intervene between the detection of features and the perception of words. Since the constituent letters provide parallel information, more information can be extracted from longer strings, leading to fast identification. Identification time might then still be equal to that of shorter words, since in longer strings lateral visual interference is higher and processing is impeded or slowed down by it.

The common problem in the pattern-unit interpretation of experimental results is that letters are denied a role in

the encoding of words. The word is encoded from the feature information in iconic memory, and features are assumed to be smaller than letters. Nevertheless, Johnson (1977) too, assumes that at some stage in the perceptual processing the features must be independent of each other, unorganized and unrelated, although they may form an integrated whole, like the set of features comprising a letter. An assembly process is the mechanism which unitizes the features into a specific encoding and for words, letter encoding is not assumed to take place. Johnson (1977) acknowledges that this interpretation of the results mentioned relies on accepting the null hypothesis. As is argued here, letter integration theories can handle the same results.
In the previous sections several results have been presented, which have been interpreted in terms of some type of pattern-unit theory, implying that higher order features, rather than letters themselves, account for word identification.
In the word superiority paradigm perceptual processing is thought to be influenced by the word context, or the meaningfulness of words. In addition to counterarguments against pattern-unit theory, it is possible to raise objections against the interpretations of the word superiority effect put forward by Reicher (1968) and Wheeler (1970). These arguments have been formulated in the context of redundancy theories. We shall discuss here the theories of Massaro, Estes and McClelland and Johnston.

REDUNDANCY THEORIES OF WORD RECOGNITION

Massaro's redundancy theory

In contrast with Reicher's interpretation (1969) of the word superiority effect, Massaro (1975a) has argued that letters are perceived in the same way in words and in meaningless strings. Massaro's arguments (1975a, b) are based on a number of experiments differing in minor but important details from the original study of Reicher (1969). Reicher (1969) implied that stimulus redundancy was completely controlled by presenting two alternative response letters, both of which could complete a word. Thus, the subject was supposed to hold sensory information in memory until the response letters appeared and then to make a decision. By this reasoning more errors would be produced when the alternatives would be similar than when they were dissimilar. This is because the detected features may fit both letters when they are similar, but usually only one when the letters are dissimilar. Thompson and Massaro (Massaro, 1975a) report an experiment in which they varied response letter similarity and found that letter recognition was not influenced by it. Consequently, letter recognition was in a way independent of the response alternatives, and of the meaningful words they formed. Apparently the perception of the letters took place before the alternatives were presented. The word stimulus, then, restricts the possibilities for its constituent letters more than either a nonword or a single letter do.

In a second experiment Thompson and Massaro (Massaro, 1975a) modified the presentation time of the response alternatives by giving them before instead of after the presentation of the stimulus. There were four alternative letters, which were learned by the subjects, consisting of two pairs of similar letters while the pairs themselves were dissimilar. In this experiment a single letter was somewhat better perceived than when it was presented in a word. Such an effect was already noted for reaction times obtained in the experiment by Johnson (1977), and was called the word inferiority effect. Massaro (1975a) argues that in these conditions the perception of single letters is no longer inferior to the perception of letters in words, because redundancy is completely controlled at the time of presentation. This was not the case with the delayed response letter presentation either in Reicher's experiment (1969) or in Wheeler's studies (1970). The conclusion is that when the response alternatives are presented, or known beforehand, redundancy is controlled and that letters in words are no longer better perceived than single letters.

Therefore the higher-order units of Wheeler (1970) are not necessary for explaining the results. That single letters are perceived in the same way as letters in words is even supported by the phenomenon that letters in words were somewhat less perceptible, which one would expect on the grounds of lateral interference. The equivalence is also supported by the effects of similarity of response letters. In word and letter conditions the errors consisted of the similar alternative in 63% of the cases, whereas 33.3% would be expected if the subjects were guessing in either of the cases. This implies that similarity has an effect only at the time the letter recognition takes place, not when the response alternatives are presented some time after the stimulus itself. The letter would then already have to be synthesized according to the theory of Massaro (1975a). Massaro (1975b) also checked recognition accuracy as a function of time. The same set of four letters had to be detected in three-letter words and nonwords where the stimulus-mask interval was varied from 10 to 240 ms. In accordance with the prediction, recognition accuracy indeed increased 40% from the shortest to the longest interval. In addition, there was no difference between accuracy for words and nonwords, again implying that letters are identically perceived, regardless of letter context.
These experiments are all concerned with recognition accuracy whereas Johnson's pattern unit model (1977) was based on latency results. Massaro and Klitzke (1977) describe an experiment on letter identification in which the dependent variable was latency and which also supports Massaro's redundancy theory (1975b). In line with Johnson's results (1977) Massaro and Klitzke (1977) found that a single test letter was more quickly identified as a target letter than a test word as a target word. This slight difference is ascribed again to the deleterious effect of visual interference. Massaro and Klitzke (1977) also present an interesting account of what would happen if a target letter had to be identified in a test word. A <u>target letter</u> could

be present in all positions of the word and would likely be equal to a letter having the same envelope. There might be several of these in the word, so a letter decision could only be made when recognition of the constituent letters had proceeded far enough, taking time. A complete target word, however, sharing superficial letter features with the test word, would very likely be equal to it, and a decision could then be made without much processing.

The results of Massaro and his associates (Massaro, 1975a, b; Massaro and Klitzke, 1977) demonstrate that letters are functional in word perception. All effects can be described on the basis of letter-bound properties and the redundancy provided by knowledge of words. This is also the central point of the letter confusion recognition model proposed in the following chapters. One factor which may be operative in the effects of redundancy, according to Massaro (1975a) is orthography. Spelling rules may allow some letters, and exclude others in given positions dependent on the perceived adjacent letters when dealing with words. Thus, if only parts of letters have been seen the decision as to letter identity is to some degree dependent on spelling context as represented in words. Letter perception as such, however, is not influenced by any context whatever.

The redundancy theory of Estes

The word and letter recognition theory of Estes is based on an extensive range of investigations but is confined mostly to the perceptual components of letter recognition in words. A basic observation in letter perception is that letters in intermediate positions of words are less well perceived than single letters at the same positions. This finding led to the interactive channels theory (Estes, 1972) which contrasted with the limited memory theory for the same results. It had been argued that instead of visual interaction, memory limitations could account for the reduced performance on longer strings. Estes (1977) denies this possibility on two grounds. First, the effect also obtains when there are only four letters to report, which is well within memory capacity. Furthermore, when eye movements are prevented, and observation times are varied from 150 to 2400 ms, performance increases slightly, but the general structure of the results remains the same. Apparently, rehearsing the letters for 2.5 seconds is not sufficient to overcome elementary perceptual limitations. In normal reading such stimulus degradation is quite common, so why is it that a reader does not notice it? Estes' answer (1977) is that readers can automatically complete words which are incompletely seen, completing them by knowledge of the words. This same interpretation is held by Bouma (1973) and is actually a basic property of the letter confusion model.

With respect to the word superiority effect Estes (1977) carried out an interesting variation of the Reicher-Wheeler paradigm. Estes (1977) notes that in Reicher's experiment (1969) both response letters would make words in the word stimulus case, and nonwords when a nonword was presented. When the response alternatives are provided beforehand, it is possible to choose stimuli in such a way that a word and a nonword can be formed.

For example, when the response alternative are A and D, the word stimulus HAT would lead to either HAT or HDT, if the context letters are seen correctly. In such a case the choice is easy. If a nonword HDT is presented the subject is not assisted by either HAT or HDT in giving his report. This set-up contrasts with that of Reicher (1969) where A and I could be alternatives for HAT, leading to HAT or HIT. For the stimulus XAG this would lead to XAG and XIG, both nonwords. Estes (1977) calls these latter conditions WW and NN respectively. The condition in which the response alternatives lead to a word and a nonword for a word stimulus is called WN, that for a nonword stimulus is called NW. Bjork and Estes (1973) applied these conditions in a letter recognition experiment, which, except for stimulus length (4 letters) and the WN and NW conditions, was the same as that of Thompson and Massaro (Massaro, 1975a). The advantage of the WN and NW conditions is that redundancy can be varied, instead of kept constant at the time presentation. In the WW and NN conditions, Bjork and Estes (1973) did not find any difference between recognition performance in words and that in nonwords, this was also the finding in experiments by Massaro (1975b). In favourable perception conditions single letters were perceived better than either words or nonwords, but in conditions in which performance decreased the accuracy level was equal for single letters, words and nonwords. Visual interference is thought to be responsible for the decreased performance on words and strings. The conclusion from these findings is again that elementary perception of letters or letter features is not influenced by linguistic context. Estes presents still other interesting evidence in support of the effects of redundancy. The arrangement of his stimulus conditions (Estes, 1977) is shown in Figure 15.

STIMULI AND MASK IN THE EXPERIMENTS

duration	exp. I	exp. II	exp. III
15-25 ms	READ	R	READ
	$$$$	$EAD	$EAD
	$$$$	$EAD	$EAD
	$$$$	$EAD	$EAD
1000-	$$$$	$EAD	$EAD
2000 ms	$$$$	$EAD	$EAD
	$$$$	$EAD	$EAD
difference between scores in conditions			
WW - NN	.0	.0	.0
WN - NW	.0	.05	.14

Figure 15

Temporal arrangement of the stimuli and masks in Estes' experiment (1977). The subject has to detect the target letter R or L. In experiment I the whole string or word is masked after ca. 20 ms by $ characters. In experiment II the context letters follow the target letter and in experiment III the context appears simultaneously with the target and remains visible.

In all three experiments the difference in probability of correct report in the WW and NN conditions is zero. The difference between the scores in the WN and NW conditions reflects the effect of redundancy and is highest with simultaneous context (exp. III). It is revealing that when the context letters appear after the target letter (exp. II) the redundancy has less effect. Estes (1977) rejects sophisticated guessing theory on the basis of this result, since in experiments II and III the subject had effectively the same opportunity to guess, which was apparently not fully employed. In addition it appeared that fast responses (< 500 ms) showed no word superiority effect, but responses taking more than 1000 ms revealed a large word advantage. This effectively rules out the notion that global features are perceived prior to letter features or letters. Fast responses involve direct matches between sensory and memory features on the level of letters, slow responses reflect the redundancy effect by other letters when a match does not succeed. In the latter case performance is also much reduced.

Estes' word recognition model (1977) has not been developed in much quantitative detail. It is hierarchical in the sense that first feature information is processed, which is fed into letter detectors, while letter information is fed into word representations in a way analogous to Morton's logogen system. The background context can also be fed into the word representations; this is done independently of the letter information. The word 'context' is used by Estes (1977) in a somewhat confusing way; word context is meant to be the letters in a word which provide redundant information for the target letter; background context is the syntactic or grammatical context surrounding the word. The advantage of the hierarchical model is that sensory information is processed independently of any context and is not subject to prior expectations of the subject. To cite Estes (1977):

> (information) from the feature-letter system (...) provides an almost fail-safe mechanism for a response to printed messages (...).

The role of global features is limited in Estes'view (1977). The main global features, word contour and word length, may only become useful when the set of alternatives is limited. A reader may identify a word correctly on the basis of global features when he has prior information on it, which, as has been said earlier, was probably the case with the results of Erdmann and Dodge (1898). Since the model of Estes (1977) is very similar to Morton's model, which is discussed later, further treatment will be omitted here.
The quantitative letter confusion model which is proposed here will also be shown to be similar to the logogen model (Morton, 1969) but it specifies the combination of letter information, which the logogen model does not. Estes has proposed a simple rule for the influence of word context on the identification of the target letter. This rule implies that detection of features occurs independently and that they are detected independently of linguistic context. Estes (1975) states, however, that the rule is somewhat arbitrary. In addition, it does not take into account interference effects between letters which are probably present in Estes' data (1977). Such interference effects would not, however, invalidate the independence implied by the rule as it is defined.

The redundancy theory of McClelland and Johnston

There is one experiment by Estes (1977) which is relevant to a study carried out by McClelland and Johnston (1977). Estes (1977) found that four-letter words were categorized much more often than chance as words when all four letters could be reported correctly. However, when less than four letters were correct, the word was categorized so much less often than chance level. This seems a somewhat trivial observation, but it provides an explanation for two results found by McClelland and Johnston (1977) concerning the interdependence of letter report in words. McClelland and Johnston presented words, single letters and pseudowords which consisted of meaningless but orthographically regular strings.

For both words and pseudowords it appeared that the likelihood of having all four letters correct was much greater than could be accounted for by independent letter identification. On the other hand, the likelihood of having three letters correct was much lower than that predicted on the basis of independence. Estes' finding leads to the conclusion that when sufficient features have been extracted from all four letter positions the word representation is probably evoked. However, it will probably not be evoked when there is information from less than the four letters. But when the word representation is evoked, linguistic information can elevate the report probability of all four letters.

In such a case independence of letter report no longer holds. Specifically, the probability that all four letters are correct should be much higher than according to an independent processing model. This is indeed found in the results of McClelland and Johnston (1977). The probability that all four letters are correct is underestimated by the independence model, but the probabilities that three, two, one or zero letters are correct are overestimated. The same situation applies to pseudowords, though there the deviations from the independent processing predictions are smaller.

However, since also in these strings letter report was more accurate than for single letter stimuli, some interdependence of reporting the letters is inevitable. When the same words and single letters were presented in another experiment, but now with unrelated strings instead of pseudowords, letter report in words showed the same deviations from independent processing as before. But letter reports in unrelated strings conformed very closely to independence predictions. Together these findings lead to the conclusion that letters are processed first in an independent fashion, but when processing has progressed so far that all letters of a word can be identified, word redundancy can complete the sensory representation successfully. It is useful to note that bigram frequency had no effect on accuracy of letter report at all, neither in words nor in pseudowords. However, word frequency did have some effect and frequency of the constituent letters in their position had an equally large effect. This rules out the possibility that word perception is mediated by identification of letter cluster units. Again, letter recognition seems to be a basic component of word perception, which can be aided by familiarity of the whole word. For their explanation McClelland and Johnston employ the recognition model of Estes (1977). In addition, they assume that when letters in words have been identified the ensuing word response can serve as an efficient code for letter report.

This explanation of the word advantage is identical with the coding-with-loss explanation given by Wheeler (1970) and the memorability explanation of Henderson (1977). To account for the better letter report in pseudowords McClelland and Johnston (1977) hypothesize that abstract, possibly phonological codes are formed, since letter-cluster units were not operative in letter report. As a case in point the Vocalic Center Groups (VCG) of Smith and Spoehr (1974) are mentioned. However, it is unclear how these VCG's might operate and how effective they really are (Henderson, 1977).

Johnston (1978) provides more evidence on the model with the results of an experiment in which contextual constraint was varied. The sequence -ATE is part of nine four-letter words, but -RIP is only part of three: DRIP, GRIP and TRIP. As in the experiment of McClelland and Johnston (1977) the subjects reported first the four letters they had seen, then chose from two similar alternative responses, e.g. DATE, GATE for the stimulus DATE. Contextual constraint did not have any effect at all. This situation can be fruitfully compared with the failure to obtain differential performance when response similarity is varied (Massaro, 1975a).

To Johnston (1978) this finding is also sufficient reason to reject Sophisticated Guessing Theory (SGT), which predicts that performance should be better when there are fewer alternatives to choose from. Johnston (1978) did obtain a word advantage when he presented the words in an experiment with unrelated strings and single letters, proving that the words were indeed redundant. Johnston (1978) discusses four models to explain the results of his and earlier studies.

1. Processing of additional features.
   Additional features might be higher order units, and a model based on them is rejected on the basis of McClelland and Johnston's results (1977).
2. More effective extraction from features in words.
   The ineffectivity of the contextual constraint (Johnston, 1978) is incompatible with a model in which it is supposed that features are extracted with more efficiency from words.
3. Sophisticated Guessing Theory.
   This theory is rejected for the same reason as the former; when sensory information has been extracted, sophisticated guessing does not apparently help, or take place.
4. More effective use of letter information in words.
   Johnston (1978) concludes that letter information can be used more effectively when they form words than when they form strings. When letter codes are lost, performance suffers more for single letters and unrelated strings than for words, where a backup word code frequently remains available.

Models 1, 2 and 4 are equivalent to the three models put forward by Wheeler (1970), so that, although the whole word superiority research has not nearly been covered in this review, the explanations certainly converge to a few common and elementary themes.

A final comment may be given on independence of letter recognition in words. Although McClelland and Johnston (1977) do not give an explanation for the deviations of letter report data from predictions according to an independence model, it is clear that there should be deviations in the case of words. Superficially, the deviations might give rise to the notion of whole word processing. However, the letter confusion model, to be presented in the next chapter, has also the property that in the predicted responses the occurrences of correct letters are interdependent. This is cuased by the selection of words, forming part of the reader's vocabulary. The probability of a word response, therefore, apparently conceals the independence of letter combination on the perceptual level.

Consequently, although the constituent letters mediate its perception, the word must be considered as a powerful linguistic unit, already on a-priori grounds. Though it is somewhat circular to say, the reader's lexicon contains words as units. But is that all there is in the lexicon? For a child learning to spell out the letters, single letters may be thought to be tiny words. When the child, trying to read a word, stumbles across a letter he does not know, he will usually not succeed in deciphering the word. But letters remain important for experienced readers as well. Also, letter perception as such seems not to be influenced by the specific context in which letters appear. So, maybe at a more elementary level of the lexical organization letters may indeed be represented as units. Sometimes they are indeed words, e.g. when we refer to an 'a' as the first letter of the alphabet. The role of letters is an intriguing one; letters are probably the first units to be extracted from the stimulus information, yet they are very close to the words we read.

REFERENCES

Ach, N. Über die Willenstätigkeit und das Denken. Göttingen: Vandenhoeck & Ruprecht, 1975.

Andrew, J. M. Why can't delinquents read? Perceptual and Motor Skills, 1978, 47, 640.

Averbach, E., & Sperling, G. Short term storage of information in vision. In: C. Cherry (Ed.) Symposium on information theory. London: Butterworth, 1961.

Bjork, E. L., & Estes, W. K. Letter identification in relation to linguistic context and masking conditions. Memory & Cognition, 1973, 1, 217-223.

de Booij, E. P. 58 Miljoen Nederlanders en de lagere school. In: A. F. Manning and M. de Vroede (Eds.). 58 Miljoen Nederlanders. Amsterdam Boek, 1977.

Boring, E. G. A history of experimental psychology. New York: Appleton Century Crofts, 1950.

Bouma, H. Interaction effects in parafoveal letter recognition. Nature, 1970, 226, 177-178.

Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. Vision Research, 1973, 13, 767-782.

Bouma, H., & de Voogd, A. On the control of eye saccades in reading. Vision Research, 1974, 14, 273-284.

Bouma, H., Legein, Ch. P., & van Rens, A. L. Visual recognition by dyslectic children. IPO Annual Progress Report, 1974, 9, 104-109.

Butler, S. R. Reading problems of Chinese children. In: L. Tarnopol and M. Tarnopol (Eds.) Reading disabilities, an international perspective. Baltimore: University Park Press, 1976.

Carmichael, L., & Dearborn, W. F. Reading and visual fatigue. Boston: Houghton Mifflin, 1947.

Chall, J. Learning to read: the great debate. New York: McGraw-Hill, 1967.

Coltheart, M. Iconic memory: a reply to professor Holding. Memory & Cognition, 1975, 3, 42-48.

Critchley, McD. The dyslexic child. London: William Heineman, 1970.

Doggett, D., & Richards, L.G. A reexamination of the effect of word length on recognition thresholds. American Journal of Psychology, 1973, 88, 583-594.

Donders, F. Over de snelheid van psychische processen. Nederlandsch Archief voor Genees- en Natuurkunde, 1869, 4, 117-145. A translation appeared in: W. G. Koster (Ed.) Attention and Performance. Amsterdam: North Holland Publishing Company, 1969.

Ebbinghaus, H., & Dürr, E. Grundzüge der Psychologie. Leipzig: Verlag von Veith & Comp., 1913.

Eisenberg, L. The epidemiology of reading retardation and a program for preventive intervention. In: J. Money (Ed.) The disabled reader. Baltimore: the John Hopkins Press, 1966.

Erdelyi, M. H. A new look at the New Look: perceptual defence and vigilance. Psychological Review, 1974, 81, 1-25.

Erdmann, B., & Dodge, R. Psychologische Untersuchungen über das Lesen auf experimentelle Grundlage. Halle am Saale: Max Niemeyer, 1898.

Estes, W. K. Interaction of background and signal variables in visual processing. Perception and Psychophysics, 1972, 12, 278-286.

Estes, W.K. On the interaction of perception and memory in reading. In: D. Laberge and S. J. Samuels (Eds.) Basic processes in reading: perception and comprehension. Hillsdale: Lawrence Erlbaum Associates, 1977.

Gibson, E. J.. & Levin, H. The psychology of reading. Cambridge (Mass.): the M.I.T. Press, 1975.

Henderson, L. Word recognition. In: N. S. Sutherland (Ed.) Tutorial Essays in Psychology Vol. I. Hillsdale: Lawrence Erlbaum Associates, 1977.

Holding, D. Sensory storage reconsidered. Memory & Cognition, 1975, 3, 31-41.

Howes, D. H., & Solomon, R. L. Visual duration threshold as a function of word-probability. Journal of Experimental Psychology, 1951, 41, 401-410.

Huey, E. B. The psychology and pedagogy of reading. The McMillan Company, 1908; Cambridge (Mass.): The M.I.T. Press, 1968.

Johnson, N. F. A pattern-unit model of word identification. In: D. Laberge and S. J. Samuels (Eds.) Basic processes in reading: perception and comprehension. Hillsdale: Lawrence Erlbaum Associates, 1977.

Johnston, J. C. A test of the sophisticated guessing theory of word perception. Cognitive Psychology, 1978, 10, 123-153.

Longley, C. (Ed.) BBC adult literacy handbook. London: The British Broadcasting Corporation, 1976.

Massaro, D. W. Primary and secondary recognition in reading. In: D. W. Massaro (Ed.) Understanding Language. New York: Academic Press, 1975a.

Massaro, D. W. Experimental psychology and information processing. Chicago: Rand McNally, 1975b.

Massaro, D. W., & Klitzke, D. Letters are functional in word identification. Memory & Cognition, 1977, 5, 292-298.

McClelland, J. L., & Johnston, J. C. The role of familiar units in perception of words and nonwords. Perception and Psychophysics, 1977, 22, 249-261.

Money, J. Dyslexia: a postconference review. In: J. Money (Ed.) Reading disability. Baltimore: the John Hopkins Press, 1967.

Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

Neisser, U. Cognitive Psychology. New York: Appleton Century Crofts, 1967.

van Nes, F. L., & Bouman, M. A. Spatial modulation transfer in the human eye. Journal of the Optical Society of America, 1967, 57, 401-406.

Patel, A. S. Spatial resolution of the human visual system. The effect of mean retinal illuminance. Journal of the Optical Society of America, 1966, 56, 689-694.

Perry, G., & Aldridge, A. The Penguin book of comics. Harmondsworth: Penguin Books Ltd., 1967.

Postman, L. Perception and learning. In: S. Koch (Ed.) Psychology: the study of a science V. New York: McGraw-Hill, 1963.

Reicher, G. M. Perceptual recognition as a function of the meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Richards, L. G. On perceptual and memory processes in word frequency effect. American Journal of Psychology, 1973, 86, 717-728.

Richards, L. G., & Heller, F. P. Recognition thresholds as a function of word length. American Journal of Psychology, 1976, 89, 455-466.

Solomon, R. L., & Howes, D. H. Word frequency, personal values and visual duration thresholds. Psychological Review, 1951, 58, 256-270.

Steinberg, D. D., & Yamada, J. Are whole-word kanji easier to learn syllable kana? Reading Research Quarterly, 1978, 14, 88-99.

Stevens, S. S. Handbook of experimental psychology. New York: Wiley, 1951.

Tarnopol, L., & Tarnopol, M. Reading and learning problems world wide. In: L. Tarnopol and M. Tarnopol (Eds.) Reading disabilities, an international perspective. Baltimore: University Park Press, 1976.

Thompson, L. J. Reading disability. Springfield, Charles C. Thomas, 1966.

Vernon, M. D. Reading and its difficulties. Cambridge: Cambridge University Press, 1971.

Wheeler, D. D. Processes in word recognition. Cognitive Psychology, 1970, 1, 59-85.

Woodworth, R. S. Experimental Psychology. New York: Henry Holt Company, 1938.

Woodworth, R. S., & Schlosberg, H. Experimental Psychology. Henry Holt and Company, 1956.

Wundt, W. Grundzüge der physiologischen Psychologie Bd. 3. Leipzig: Wilhelm Engelman, 1903.

Wundt, W. Völkerpsychologie Bd. 1. Leipzig: Wilhelm Engelman, 1911a.

Wundt, W. Vorlesungen über Menschen- und Tierseele. Hamburg: Leopold Voss, 1911b.

Wundt, W. Grundriss der Psychologie. Leipzig: Alfred Kroner Verlag, 1914.

2 A Model for the Visual Recognition of Words of Three Letters

Don Bouwhuis
Institute for Perception Research
Eindhoven, The Netherlands

reprinted from. ATTENTION AND PERFORMANCE VII
Proceedings of the Seventh International Symposium
on Attention and Performance, Sénanque, France, August 1-6, 1977
edited by J. Requin; Hillsdale. Lawrence Erlbaum Associates, 1978.

The model to be presented predicts recognition performance in a situation where subjects respond to Dutch words consisting of three letters presented at different eccentricities in the visual field

As basic features of the word stimulus the constituent letters were chosen, being about midway between elementary visual properties and global word-shape This choice is empirically supported by a strong relation between letter recognition in words and in meaningless strings

The perception of letters in such a case is taken to be position-specific, as letters in words are subject to interaction by neighboring letters dependent on eccentricity By means of a probabilistic rule, letter recognition, given a stimulus word, leads to a set of letter strings as viable alternatives From these strings only the real words are retained by means of a matching procedure for which the final response probabilities are predicted with the Constant Ratio Rule Thus word frequency effects are not separately incorporated in the model

All parameters of the model are supplied by earlier results on letter recognition Though the reliability of these data is not optimal, predictions of the model compare favorably with responses obtained in a word recognition experiment Tests of the model are described for correct and incorrect responses to words presented at four different eccentric positions

I. INTRODUCTION

If one considers that reading involves word recognition, and word recognition involves recognition of simpler elements, word recognition is therefore important for both the study of reading and the study of elementary recognition Emphasis on the latter approach leads to experiments with more tightly controlled conditions than are possible in a reading situation and, of course, need not be concerned with words at all. Even if words are studied, the aspects under investigation may require the elimination of the typical reading setting. Eliminating the conditions favoring correct recognition one by one reveals more clearly the properties of the recognition function but can in itself not produce a clear insight in what happens in reading.

In this paper a word recognition paradigm was modified to more closely resemble a normal reading situation. However, in order to have adequate experimental control over the stimuli, the stimuli used were single words. This situation must, of course, somewhat depart from reading. Nevertheless, the present approach accounts for one important step in reading: the perception of words from a string of letters somewhere within a line. Specifically, responses to briefly presented words were predicted by means of letter recognition data and word knowledge. Much is already known about letter recognition, by including word knowledge a model has been developed that predicts not just changes in word response probabilities as a function of conditions, but these probabilities themselves. The present approach, therefore, contrasts with studies on factors influencing recognition accuracy or recognition time.

In the next section stimulus aspects used in recognition studies will be compared with those in reading studies. The model itself will then be presented, followed by a test using data from experiments by Bouma (1973). In the discussion, some properties and limitations, as well as possible extensions of the model, will be elaborated.

II. READING VERSUS RECOGNITION

A. Word Contour

In the early days of experimental psychology, research on reading was more intensive than it has been during the first half of the twentieth century. Starting around 1880, most research workers then employed fragments of actual text. Huey (1908/1968) worked with lines cut from the American Journal of Psychology and was able to demonstrate a number of phenomena of prime importance with this material. The quality of his observations, which have been corroborated in recent times, makes up for his lack of statistical evidence. Experimental deviation from reading situations started, however, quite soon afterward. Snellen, who was the successor of Donders, introduced a visual screening test, consisting of single capital letters with progressively reduced size.

Now small size has not found wide application in recognition studies, but single presentation either as letter, letter string, or word has. Likewise, the use of capital letters has been pervasive in word recognition experiments, although it is not entirely clear why. The word contour is eliminated, but though vertical letter size is standardized, letter width is not, which is only partially compensated by equal spacing in some type faces Similarity is not controlled either, but in any case perception is more difficult, because "dominant parts," as Huey (1968) called them, are eliminated Dominant parts are provided, among other ways, by word contour, as determined by the succession of short letters, ascenders and descenders (Table 1)

The function of word contour is one of the basic issues in the early investigations on reading It was supposed, by Cattell, Erdmann, and Dodge, Dearborn, and others, and still is, that the perception of words is mediated more by their total outline than by their constituent letters A more detailed view was expressed by Zeitler, Goldscheider and Muller, and Messmer (Huey, 1968) stating that dominant parts of words alone could trigger their recognition Huey (1968) noted that "Total form is not perceived separately, but that, in one act of projection, the total form and the parts to fill it are placed " This scheme closely resembles the Gestalt view to be introduced some twenty years later, which has only rarely, if ever, resorted to words as illustrations for the unitary forces in perception

It is true that word perception undoubtedly entails more acquired components than figure ground segregation, but this does not dispel the notion of Gestaltlike properties of words In order to experimentally impair word perception it should therefore be quite effective to obliterate word contour While printing in capital letters just standardizes contour, a reader should be expected to have even more trouble with modified contours These have been produced by mixing upper and lower case letters within a word

Another drastically altered word contour is arrived at by printing the letters vertically aligned, apart from being unfamiliar for the reader, it forces him to change scanning direction Without any formal theory related to the perception of the constituent parts of words, detailed predictions of recognition responses with these procedures are beyond reach

similarity groups of letters

short letters	a s z x
	e o c
	n m u
	r v w
ascenders	t i l f
	d h k b
descenders	g p j y q

TABLE 1 Classification of letter types Similarity has been established in a range of recognition experiments

B Luminance

The most straightforward way to reduce visibility is to decrease luminance For an effective degradation, however, the luminance level must be much lower than 30 cd/m^2, below which the contrast sensitivity is also impaired, leading to reduced resolving power Therefore, a luminance level exceeding 30 cd/m^2 is generally acceptable for reading, but reduced luminance essentially requires stimulus size increments to compensate for the lower contrast sensitivity

The highest liminance that is usually reported in the literature seems to be about 10 fL, corresponding to 34 cd/m^2 Characters are usually presented at lower levels, and thus are viewed under lighting conditions that are suboptimal for reading

C Duration

Increases in stimulus duration beyond 30 msec seem to increase word recognizability very little provided poststimulus masking does not occur within 100–200 milliseconds from stimulus offset

Normally we find eye fixations in reading pauses to be minimally of the order of 100 msec, with an average of 200 msec (Andriessen & de Voogd, 1973) It is not clear, however, whether the processing of the visual mechanism going on in the extremely short presentations of 2–3 msec sometimes applied, can be directly related to that during the much longer regular eye fixation pauses

D Type Font

In order to be sure that the visual attributes of the stimulus presented for short durations can be optimally processed, the word or letters are often made up in a large and simple type face This is especially the case in so called lexical decision tasks in which no visual errors or confusions are allowed In both tasks, recognition and decision, the word or letter string extends over a large area of the fovea by which visual interference relative to a smaller word in a regular typefont is reduced

III EYE POSITION AND VISUAL INTERFERENCE

As was noted already, single stimulus presentation is practically the rule in recognition tasks that do not imply search In printed text the reader is scanning one line at a time, which is completely filled with words, whereas there are nonattended lines above and below the one under consideration With eye position fixed, only the small portion of the letters or words surrounding the fixation point can be perceived clearly Huey (1968) cites results of Erdmann and Dodge, and of Cattell, who found that from a short eye fixation on a line of connected text, a maximum of about five words could be reported. The words ranged in length from two to ten letters with five letters on the average. However, only four or five letters could be reported when letters were presented in a random arrangement, which is of course the same result as that found by Sperling (1960). Sperling was also able to show that, during a short time after presentation, more letters were actually available than were reported. This might partially explain the fact that, when the stimuli were unconnected words, two words could be reported (Cattell, cited by Woodworth & Schlosberg, 1956). This is twice the number of reportable unconnected letters; apparently the subject is able to organize the available letters into word structures under these conditions. However, connected words are even better recognized, and in this case it is clear that not all details could have been perceived. Figure 1, derived from data of an experiment by Bouma (1973), shows the probability of correct report of words averaging four to five letters as a function of eccentricity.

At eccentricities exceeding 1° visual angle the probability of correct recognition decreases sharply. This result is only valid for singly presented words. Assuming that if there are more words in the display, they do not influence the perception of each other (which is not true), the average number of words having four–five letters that can be perceived in one line in one eye fixation is about five. This number can be estimated from the curve in Fig. 1 yielding three words on the right side and two on the left. Visual interference, however, would limit this number considerably but apparently context makes up for that part of a meaningful line that has not been perceived sufficiently well.

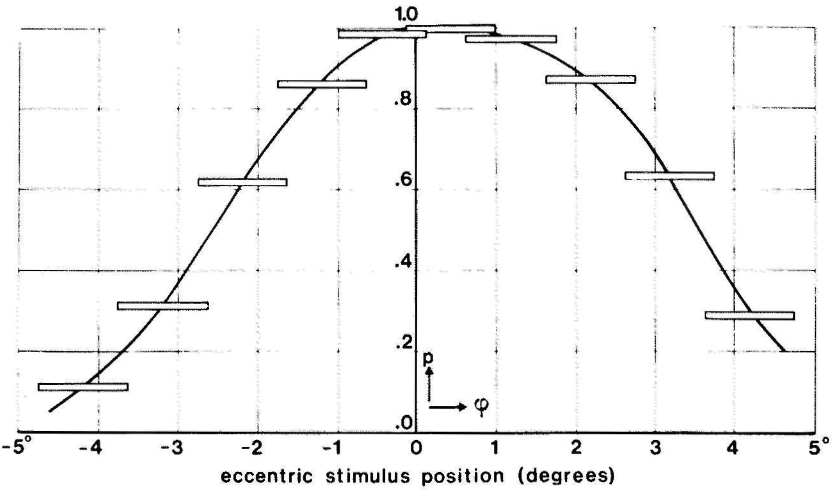


FIG. 1 Recognition scores of words, ranging in length from 3-6 letters (average 4.5 letters) as a function of eccentricity of presentation for 100 ms. Note the right-field superiority. One degree comprises four letters. Data from Bouma (1973).

The relevance of the eccentric position of words during reading is made clear by the fact that eye saccades average about eight letter positions, whereas 1° visual angle comprises four letter spaces. In as much as the extent distribution of left to right saccadic movements is positively skewed, much larger jumps of up to 20 letters (5° visual angle) can occur. For an account of the visual interference effects of this saccade based eccentricity, the reader is referred to the chapter by Bouma on the functional visual field in this volume.

At this point we may elaborate somewhat more fully on the concept of global wordshape. Aside from being vague, global wordshape is an unstable property of a word, because of the widely varying interference over the visual field, which will continuously modify it.

Wordshape is conveyed by the constituent letters of the word. Figure 2 shows the perception of separate letters, initial and final ones only, in words and meaningless strings. Apparently there is a close correspondence between letter perception in both cases, though performance is better on words. So the problem of global wordshape could possibly be restated in terms of the contribution of the constituent letters and their mutual interference effects

That letters are more accurately reported in words may be based on the fact that readers do not need to perceive all letters correctly in order to recognize the word, and therefore may sometimes be able to infer a letter that has not actually been seen. This of course cannot hold for strings, which is why the superior performance in words has been called completion (Bouma, 1973). Word recognition therefore minimally requires letter perception and word knowledge. Both factors have been incorporated in the model to be presented.

IV. THE RECOGNITION MODEL

We will take the situation in which the subject tries to recognize real words which he knows to have three letters. It is supposed that he perceives a letter for any of the three positions of the word, that is, the internal representation is a string of three letters Which letters will be perceived depends on the position in the word and on the eccentricity of the word It should be noted that the dependence is not stated in terms of the adjacent letters, but just taken to be position-specific In this way interference is in a sense a stable property within words For any position in the word, a confusion probability can be defined that a letter belonging to the stimulus word, l_i, is perceived as λ_j, which can be written for the middle position as

$$P_{\phi}(-\lambda_j - \mid - l_i -) \tag{1}$$

The subscript ϕ denotes the eccentricity of the stimulus word. Clearly, probability 1 defines a confusion matrix. Furthermore, the simplest way in which these

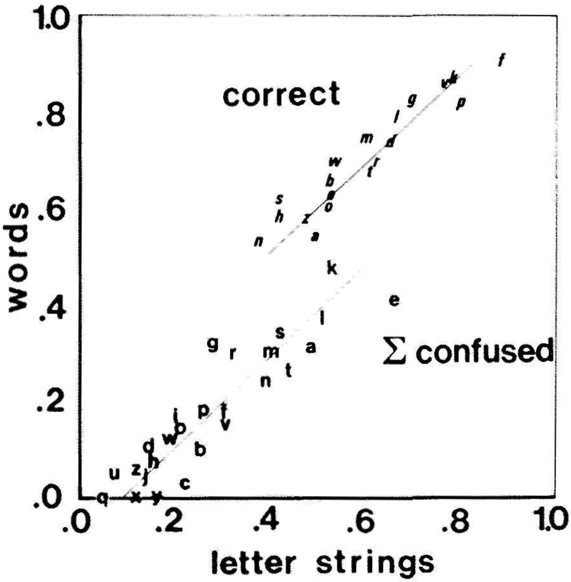


FIG. 2 Comparison of letter scores in words and meaningless strings. Upper right data points represent fractions correct. The lower left points represent the fractions of incorrectly reported letters: the response probability minus the fraction correct. Data from Bouma (1973).

probabilities can be combined is by taking them independent of each other. Then the probability of the activation of any string $\lambda_h\lambda_i\lambda_j$ can be written as:

$$P_\phi(\lambda_h\lambda_i\lambda_j|l_pl_ql_r) = P_\phi(\lambda_h--|l_p--)\,P_\phi(-\lambda_i-|-l_q-)\,P_\phi(--\lambda_j|--l_r) \quad (2)$$

For one eccentricity ϕ the probability of a letter string perceived for the presented stimulus word can be obtained by multiplying the appropriate terms of three confusion matrices, one for each letter position. Other ways of combination are conceiveable, and generally desirable, and a consideration of some aspects of the independence assumption will be postponed until the discussion.

The second stage in the model represents the effect of word knowledge. It has been assumed that on every trial a large number of strings is activated; most of them will be meaningless, but some will be real words. It appears, however, that observers faced with the task of recognizing words only rarely report nonsense strings. When not pressed for an answer, they may sometimes state not to have seen the stimulus at all. For the final choice of the word response, then, it is supposed that all meaningless strings are rejected, whereas all activated words are retained. With regard to the reported words, it is generally found that visual factors are much more important than effects owing to frequency of occurrence of words (Morton, 1969, Rumelhart & Siple, 1974). Since in normal reading context as well as visual factors are operative, the role of word frequency would be expected to be even smaller there So the basic issue here is how much of word recognition can be accounted for by visual factors and word knowledge. Consequently, in the model the response is chosen from among the real words activated by the letters of the stimulus word according to the Constant Ratio Rule (Clarke, 1957). This can be formally expressed as·

$$P(r_k \mid w_s) = \frac{P(r_k)}{\sum_l P(r_l)} \tag{3}$$

where the r_l are the words among the activated strings $\lambda_h \lambda_i \lambda_j$, and w_s is the presented stimulus word corresponding to $l_p l_q l_r$ in probability 2.

In this way, the word recognition process is completely covered in a minimal fashion, by activation of letter strings dependent on the stimulus, selection of words, and the prediction of response probabilities by the Constant Ratio Rule. In the form the model is stated there are some mathematical correspondences with the earlier formulations of Morton (1969) and of Rumelhart and Siple (1974). However, in Morton's model there is no account of the processes going on before the triggering of the logogens or word alternatives In Rumelhart and Siple's model (1974), letter confusions are predicted theoretically and there is no letter position effect. additionally, the choice rule works already on the level of individual letters.

V. AN EXPERIMENTAL TEST

For a test of the model two different kinds of data are needed. First, letter confusions that occur for letters in different positions of words themselves, presented at various eccentricities. Second, the set of all known Dutch words of three letters is required to specify which of the perceived letter strings are words.

Experimental data relevant to the letter confusions were collected by Bouma (1973). His experiments were directed at the contribution of initial and final letters of words to their recognition in the eccentric visual field. In order to obtain pure estimates of the recognition of these letters, they could not be presented in words because letter completion would occur. Words were therefore transformed into unpronounceable strings by exchanging all letters not asked for with visually similar letters as established by Bouma (1971). In this way the visual characteristics of the words were maintained as closely as possible (Table 2). Subjects reported the initial and final letter of the string presented randomly to the right or the left of the fixation point. In this way letter confusions of the first and last letters of the three-letter words used by Bouma (1973)

WORD TRANSFORMATION

word		string	report
gas	→	gzs	initial and final letters
lip	→	lfp	
arm	→	sru	middle letter
fee	→	ieo	

TABLE 2 Illustrations of how words are changed into meaningless strings to be employed for letter recognition For the recognition of the middle letter in /arm/ the letters *a* and *m* are replaced by similar letters (Table 1) in order to obtain a visually similar string The subject is then asked to report the middle letter

were obtained In order to obtain letter confusions for the middle letters, a supplementary experiment was run, in which the words were again transformed with respect to the first and last letter The words themselves were also presented in a recognition task, in which the subjects reported whole words The model should predict these word responses In both the original experiment of Bouma (1973) and the supplementary one, there were 100 words of three letters, each one presented twice, once right, and once left of the fixation point, for 100 msec at a background luminance of 100 cd/m^2 For the typeface see Table 2 The 100 words were split up in two groups for presentation eccentricities ±3° and ±2° The same eleven subjects took part in all experiments, for further technical details the reader should consult Bouma (1973) For the second stage of the model, which involves a selection of words from activated strings, a word vocabulary was needed This vocabulary was composed of words appearing in two word counts De la Court, as published by Linschoten (1963) and Uit den Boogaart (1975), making a total of 409 words of three letters Later the list was found to be incomplete and a number of commonly known words from several sources were added to it, extending it to 541 entries It was established experimentally afterwards that words that did not appear in this list are not generally known.

VI RESULTS

A Correct Responses

Figure 3 shows the predicted and experimentally obtained proportions of words correctly identified at the four eccentricities employed One set of predictions was made for the 409-word list, another for the 541-word list Both sets of predictions follow closely the experimental values, but there is an underestimate This amounts to 3% for the 409 entry list and 8% for the 541 entry list Whatever

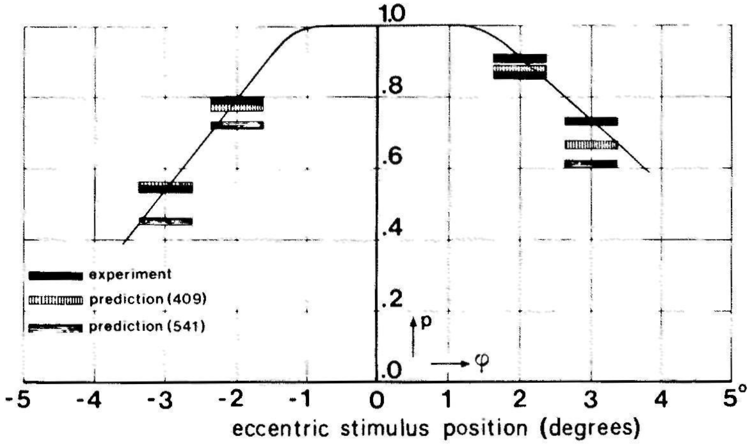


FIG. 3 Experimental and predicted probabilities of correct word recognition as a function of eccentricity. Two sets of predictions by the model are shown: one with a list of 409 words, the other with a list of 541 words.

other causes there may be for the underestimation, this difference is to be expected on the basis of the applied Constant Ratio Rule. When more word alternatives exist for a stimulus word, the probability of the correct response will decrease relative to a situation with fewer alternatives. An important point to notice is the superior right field performance, which is also present in the predictions. It has been assumed that the left hemisphere of the brain, corresponding to the right visual field, is engaged specifically in the processing of language material.

Apparently, since the model is based on the recognition of constituent letters, the specialization might well be limited to letters or even less complex features of the word. An explanation in terms of less visual interference in the right field is definitely simpler than language specialization on the level of words.

B. Letter Recognition and Completion

In addition to correct responses, the model predicts incorrect responses. An incorrect word may still have one or two letters in common with the stimulus word. This is relevant to the present model, if the letters appear in the same place as in the stimulus word. Thus an analysis of the correctness of letters in the responses includes both the correct and the greater part of the incorrect responses. In Fig. 4 the proportions of correct letters predicted by the model are compared with the proportions appearing in the responses of the subjects. Proportions are averaged over the letters appearing in the stimulus words. Also shown are the probabilities that these letters are correctly recognized in their position

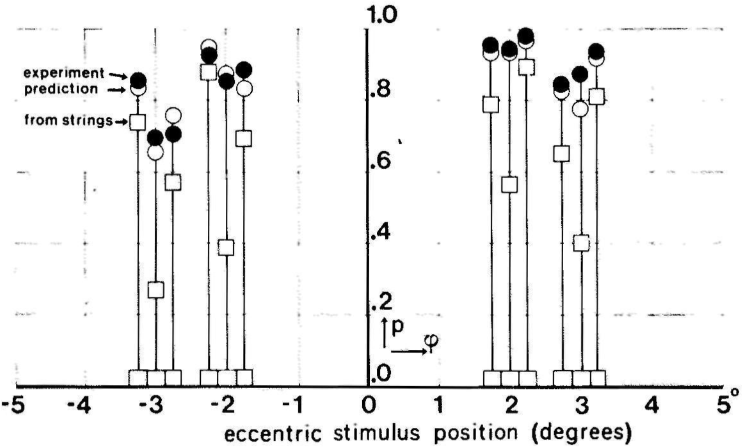


FIG. 4 Experimental and predicted probabilities of correct letter recognition in the words presented as a function of position in the word and eccentricity. The lower squares represent averages of letter scores in meaningless strings as used in the model.

in an unpronounceable string. Completion is defined as the difference between the letter recognition probability in words and the one in meaningless strings. Evidently, most completion takes place in the middle positions of words where letter recognition is weak.

VII. DISCUSSION

A model is proposed predicting single word recognition and letter recognition under conditions relevant for reading situations. The only factors taken into account were the recognition of letters in specific word positions, combined with an independence assumption, and word knowledge. The results of a test of the model have provided evidence that a substantial part of word recognition under the conditions stated can be explained without reference to general word shape, dominant parts, or frequency effects. Actually, the concept of general or global word shape has been shown to be amenable to specification because its properties can be explained in terms of interference-dependent letter recognition. Recognition of letters flanked on both sides by other letters is severely impaired in the eccentric visual field which results in a word shape different from that on fixation. Word knowledge in the observer enables him to complete the missed features of the word, leading to apparently good letter scores. Though letter recognizability in meaningless strings is much lower than that in words, the model can successfully predict scores for letters in words from that in strings.

A. Word Frequency

The experiments reported by Goldiamond and Hawkins (1958) showed that word frequency may influence responses under quite difficult perception conditions. Recent experiments by Richards (1973a, 1973b) replicating the original Goldiamond and Hawkins experiments demonstrated that the frequency effect is strongly influenced by the demand characteristic of the task It is very improbable that the demand characteristics stimulating the frequency effect were present in the experiments, considering the recognizability of the words and the lenient instructions Neither would reading seem to be so demanding Apart from this, frequency data could only have been used in a very crude way, since for at least one-third of the Dutch three-letter words, frequency of occurrence is unknown, being less than 10^{-6}

B Interaction and Independence

It is often mistakenly thought that mutually interacting processes should also be stochastically dependent Interacting processes occur in the perception of a *number of adjacent letters, where the presence of one letter interferes with the* perception of another In such a situation the probability of perceiving a letter may still be independent of the probability of perceiving one of the neighboring letters. This may be true even though the recognition probabilities may be decreased relative to a situation without interference by other letters. Imposing this independence condition on letter perception considerably simplifies the recognition theory, because there is no need to estimate probabilities of parameters specifying the dependence relation.

The present approach was intended to study how much of word recognition can be accounted for with such an independence assumption There are, however, also practical reasons for such an approach. Experimental support for dependency parameters must be unreliable in view of the relatively small number of observations on all possible different letters, letter positions, and positions in the visual field.

Some sources of letter recognition dependence will now be considered in detail Without loss of generality, the effect of dependence can be shown here by considering two letters of the word

Let P_1 be the probability that a letter, not necessarily the correct one, has been perceived for a presented letter, and P_2 an analogous probability for an adjacent letter. The covariance expresses the degree of dependence between these two variables as follows

$$Cov\, P_1P_2 = EP_1P_2 - EP_1 \quad EP_2$$

From this it follows immediately

$$EP_1P_2 = EP_1 \quad EP_2 + Cov\, P_1P_2 \tag{4}$$

Except for the covariance term, equation 4 is analogous to the letter combination rule 2 of the present model. The left term gives the letter perception in words, which is predicted by the product on the right, implying the independence condition, whereas the covariance is assumed to be near zero. However, when there is a positive covariance between the recognition probabilities, the independence based prediction will underestimate the probability of correct word recognition. This could be a second explanation of the underestimation of the correct responses by the model.

Letter recognition will covary when the subject looks to the left or right of the fixation mark during presentation. The letters on the fixated side will be perceived more clearly than those on the other side. Considering the fairly low recognizability of some letters and the absence of visually unexplainable guesses, the occurrence of this behavior would seem improbable. In general, observers – certainly when they are experienced – fixate quite accurately. It can further be shown that positive covariance can be induced by variation in the recognition probabilities (Bouwhuis, 1973). One source of variation is that observers differ in recognition performance. Since recognition performance is distributed over several letters and letter positions, the small number of comparable observations would lead to unreliable individual predictions. Another source could be that some letters interfere more with the perception of adjacent letters than others. The number of possible combinations of three letters, or even two for that matter, renders the experimental study of differential interference unrealistic.

Finally, negative covariance would arise if subjects attend to only one letter position at the expense of others. Since subjects can report up to four or five letters from a larger display, and even more in words or text, this strategy would seem improbable in the experiments under consideration. Experiments are in progress to obviate some of these limitations, though apparently not all of them can be overcome simultaneously. Theoretical developments could be useful too, because an explicit formulation of interference effects over the visual field and within words would greatly reduce experimental effort connected with the letter confusions. Contribution to the close fit of the model by any of these factors could of course still only be moderate, but might be required in a more accurate prediction for individual words.

C. Extension of the Model

As regards developments of the model, the first extension seems to lie in the direction of longer words. This will imply that the model itself must be modified with respect to middle letters. It is hardly feasible for subjects to report, or even see, middle letters of longer words in the eccentric field. Besides, length confusions tend to occur in these cases, which cannot be accounted for by the present position-based approach.

Since the number of words in those cases is appreciably larger, it is questionable whether the Constant Ratio Rule will suffice any longer More complicated choice theories will then probably be required As a first step, this very simple model has the advantage of needing no free parameters, which probably cannot be avoided in more rigid formulations for more complicated perception processes The simplicity of the present version of the model indicates that it could be profitable to test formal recognition models in reading situations

REFERENCES

Andriessen, J J , & de Voogd, A H Analysis of eye movement patterns in silent reading *IPO Annual Progress Report* 1973, *8*, 29–34

Bouma, H Visual recognition of isolated lower-case letters *Vision Research* 1971, *11*, 459–474

Bouma, H Visual interference in the parafoveal recognition of initial and final letters of words *Vision Research*, 1973, *13*, 767–782

Bouwhuis, D *Het Klemtoonoordeel, persoonlijk of eenstemmig?* 1973, IPO Report no 253

Clarke, F R Constant ratio rule for confusion matrices in speech communication *Journal of the Acoustical Society of America*, 1957, *29*, 715–720

Goldiamond, I , & Hawkins, W F Vexierversuch The log relationship between word-frequency and recognition obtained in the absence of stimulus words *Journal of Experimental Psychology*, 1958, *56*, 457–463

Huey, E M *The Psychology and Pedagogy of reading* Cambridge The MIT Press, 1968 (Originally published , 1908)

Linschoten, J *De la Court's frequentietelling van Nederlandse woorden* Utrecht Psychological Laboratory, Rep no 6301, 1963

Morton, J Interaction of information in word recognition *Psychological Review*, 1969, *76* 165–178

Richards, L G "Vexierversuch" revisited A reexamination of Goldiamond and Hawkins' experiment *American Journal of Psychology*, 1973, *86*, 707–715 (a)

Richards, L G On perceptual and memory processes in the word frequency effect *American Journal of Psychology*, 1973, *86* 717–728 (b)

Rumelhart, D E , & Siple, P Process of recognizing tachistoscopically presented words *Psychological Review*, 1974, *81*, 99–118

Sperling, G The information available in brief visual presentations *Psychological Monographs*, 1960, *74*, no 11, Whole No 498

Uit den Boogaart, P C *Woordfrequenties in gesproken en geschreven Nederlands* Utrecht Oosthoek/Scheltema en Holkema, 1975

Woodworth, R S , & Schlosberg, H *Experimental psychology* New York Henry Holt and Company, 1956

3 Word Knowledge and Letter Recognition as Determinants of Word Recognition

The recognition of words during reading does not only imply that letters, or other constituent parts of words are seen, but also that words are known. In a recognition model for words of three letters, developed by Bouma and Bouwhuis (1975) word perception is thought to be mediated by recognition of the constituent letters in their position and by knowledge of words. It was supposed for reasons of simplicity that subjects knew the same words, that the words were known equally well and that frequency of occurrence of words did not influence response probability.
In the experiment reported here these assumption were put to test by presenting all Dutch words of three letters to 30 subjects and asking them if they knew them. Perceptually similar nonwords were also presented, making the task a lexical decision task. The most important result is that words vary considerably in familiarity, which appears both from correct responses and response times. It appears too that word knowledge is rather uniform over the 30 subjects. Frequency of occurrence of words had only a minor effect; involving only the most rare, often unknown words. It is argued how incorporation of the differential word familiarity in the vocabulary of a model could improve its predictions of word recognition.

When we read printed lines of text new words keep moving into our field of vision; at the same time others are being decoded and recognized and attached a meaning. During this process expectations are formed concerning the words to come. The reader is usually able to guess words that might logically follow those he has just read. It is only by his knowledge of words that he can think of words that are suitable in a context. An interesting point about word knowledge is that a word which does not form part of it, is not a word for that reader, while it may be perfectly acceptable to readers who know it. It is not difficult to find such words and this observation merely stresses the

Also published in P.A. Kolers, M. Wrolstad and H. Bouma, Processing of visible language I, New York: Plenum Publishing Corporation, 1979, p.269-281.

fact that words are so overwhelmingly numerous that they cannot all be known by any individual. Words in a language vary from the common words of the vernacular to very rare instances of specialized jargon

WORD RECOGNITION

The recognition of words involves both perceptual and linguistic aspects, the latter being brought into operation by knowledge of the language. A general description of word recognition should, therefore, include a specification of word knowledge. The first problem that arises, then, is that one reader may know some words which another reader does not know; consequently, word knowledge would have, next to general components, substantial individual components.
For words the 'common core' principle may be expected to hold, implying that a substantial part of words is generally known. It is probably possible to assess the words individually known by a subject, but it is not quite realistic.
A description of word recognition is then only feasible and productive when people have a major proportion of their word knowledge in common. This may still imply the case that there are words which are well known and another set which are known scarcely or not at all.

WORD COUNTS

It is not uncommon to use word counts as an estimate of the words which are known. Word counts have been published by Thorndike and Lorge (1944) and by Kučera and Francis (1967) for American English and by De la Court (Linschoten, 1963) and by Uit den Boogaart (1975) for Dutch. All these lists have been compiled from printed material in which some words occur frequently, others occasionally and some not at all. Several studies have shown that words which occur frequently are identified somewhat better (Broadbent,1967). A possible explanation which comes to mind is that words which are regularly encountered by the reader are processed more efficiently than others. The frequency of occurrence of words might thus reflect how familiar they are to readers. Though it is not sure which words a particular reader has seen and how often he has seen them, word frequency lists might be considered to represent a rather general body of word knowledge, but one in which fairly uncommon words, though perhaps widely known, are usually not represented.

Rumelhart and Siple (1974) used the three-letter words in the list of Kučera and Francis (1967) for their recognition model for words of three letters. Word frequency was applied in such a way that infrequent words had a lower predicted probability of being responded than frequent ones. The fact that their model was somewhat complicated and that their experiment took in so many words and strings (726), prevented the parameters of the model from being accurately estimated, thus precluding a clear idea of how important the incorporation of word frequency in the model was for the actual predictions. Nevertheless, examination of the experimental data suggests that frequency effects are small compared to the ever-present effects owing to visual aspects of the words presented.

In all, there appear to be two ways in which word frequency counts fall short of describing word knowledge. First, variations in the knowledge of words as expressed by their frequency of occurrence has no basic impact on word recognition. Second, word counts include few of the most uncommon words, of which there are quite a large number. This fact was encountered during the development of the recognition model by Bouma and Bouwhuis (Bouma & Bouwhuis, 1975; Bouwhuis,1978), a model also intended to describe recognition of words of three letters.

I intend to describe first the way in which the word knowledge of the reader was implemented in the model. The word vocabularies adopted for testing the model influenced the model predictions in a specific manner which will be described below. The shortcomings of the vocabularies used provided the motivation for investigating knowledge of words of three letters. As will be seen, the results shed light on the general situation with regard to word knowledge among readers and indeed indicates what the most important aspect of words is, namely whether they are known or not.

SHAPING A VOCABULARY

The word recognition model of Bouma and Bouwhuis (Bouwhuis, 1978) describes the perception of a word as the process by which the constituent letters are perceived in their position. Perception is imperfect due to parafoveal presentation which causes visual interference to impede letter perception. Thus, a presented word may give rise to possibly a large number of strings of three letters, perceptually similar to the word. Some of these strings form words and it is assumed that in his response the reader takes only those words into consideration. The decision rule which is assumed to operate is the constant ratio rule (Clarke,1957); essentially its application means that the probabilities that the word forms are present among the possibly perceived strings are normalized to add up to 1.0. If such a model is to describe word recognition properly, it must employ the same lexicon as the general reader. In accordance with the principles oulined above, the appropriate words were taken from the most recent word count, that of Uit den Boogaart (1975). His list was compiled from 720,000 words of printed text and contained 1651 three-character sequences, including a number of actual words. All character sequences between two spaces were included in the count, so that the list also contained abbreviations (TUC, KLM, CIA), coding symbols (CO2, DC8, SO2), plural forms (A'S, B'S), numbers (3.6, 100, 010) and match scores (1-0, 3-4). Though these forms are interesting in their own right, they were not considered to be words and not admitted to the vocabulary used for the model. The list also contained homographs, different words spelled the same. Some are lexically different and some are produced by verb conjugations (<u>arm</u>: body part, poor; <u>was</u>: (I,he)was, laundry, wash, growth, wax).

Since words are presented in isolation in the word recognition experiment no intended specific meaning can be inferred from any context. These identical spelling forms need, therefore, only appear once in the vocabulary of the model. The assumption underlying this decision is that different meanings do not compete for recognition under the same spelling form during the recognition process, but that only one is functional at any one time, which is experimentally supported by Forster and Bednall (1976). Without going into detail it may be noted here that homographs are somewhat more quickly identified than unique spelling forms (Rubenstein, Garfield & Millikan, 1970; Forster & Bednall, 1976)[1].
Thus revised, the list of Uit den Boogaart (1975) comprised 409 different three-letter words and was used directly as the word vocabulary of the model. It was soon obvious that this list was incomplete; earlier in the recognition experiment by Bouma (1973), for which the recognition model predictions were intended, some words were reported that did not appear in the vocabulary adopted. It was, therefore, supplemented with the few words from Linschoten's compilation (1963) of the pre-war De la Court count which did not appear in the list of Uit den Boogaart (1975). De la Court had counted some 1,000,000 words and among them there were 329 different three-letter words. Finally, the vocabulary was enlarged to include the results of questionnaires concerning three-letter words, distributed among a group of readers to ensure that the words were known. Altogether these operations resulted in a total of 541 words. In principle, the predictive power of the model should have been increased by this expansion since more responses could be accounted for.

RECOGNITION AND THE ROLE OF KNOWLEDGE

The recognition model was tested with both vocabularies, the incomplete one with 409 entries and the supplemented one comprising 541 entries, on the data originally

[1] this suggests that the human lexicon may indeed contain multiple representations of homographic words.

collected by Bouma (1973). First, as was to be expected, the model's predictions of erroneous reponses were more accurate with the larger vocabulary. Second, the probability of correct responses was underestimated; by 3% in the case of the smaller vocabulary but by 8% in that of the larger one. This underestimation, especially with the large vocabulary, can be explained by the nature of the decision rule applied. When there are more possible candidate words for a presented word, the normalization of the constant ratio rule has the effect of making the individual word response probabilities lower than when there are fewer candidate words. There are simply more competitors for the response to a word and this affects the correct responses most, since they are usually the largest. From these considerations it appears that both vocabularies have inadequacies: the smaller one lacks accuracy compared with the larger, but the larger leads to greater underestimation of the predictions of correct responses. A reconsideration of the all-or-none character of the fixed vocabulary seems to be called for.

What are the consequences for the recognition model of a vocabulary in which the words are not equally well known? If the additional words in the large vocabulary were less familiar in some way, they could, when activated by a stimulus word, compete less with better known alternatives. After normalization of the word probabilities, the little known words would have a lower probability of ever producing a response, and, consequently, the well-known words would be more likely to produce a response. Since the stimulus words in the study by Bouma (1973) were well known, their predicted probabilities for correct identification would increase - precisely the effect aimed at.

Even the smaller vocabulary might contain little known words. If this factor could somehow be incorporated in the vocabulary the slight underestimation with the smaller vocabulary would also be offset. Nonetheless, the larger vocabulary would appear preferable since it allows predictions for a greater number of words and the responses to them.

It would even be possible to increase the number of words of the vocabulary for greater generality. Since additions would probably be even less well known than the words adopted, they would, consequently, have minor effects on the response probabilities of the latter.
There is ample reason to study the degree to which words are known, a task that is feasible in the case of three-letter words since these are limited in number. That words vary in familiarity seems only logical. It is nevertheless surprising to see to what extent word recognition can be described on the assumption that words are equally well known (Bouma & Bouwhuis, 1975).

EXPERIMENT

In contradistinction to a word recognition test, words presented in a word-knowledge test should be perfectly readable since the operation of word knowledge must not be hampered by visual degradation. Since a subject can always pretend to know the words presented, nonwords must also be presented, so that the test becomes a lexical decision test. In other words, the subject decides whether the string of letters presented forms a word or not. This type of task is not uncommon (Rubenstein, Garfield & Millikan, 1970; Rubenstein, Lewis & Rubenstein, 1971), but it rarely involves more than 200 words, and never a complete set like the present set of Dutch three-letter words.

Material

All Dutch words consisting of three letters were taken from the largest up-to-date Dutch dictionary (Kruyskamp, 1970). Conjugated verb forms did not appear as such and were derived. In all, 713 three-letter words were found. Meaningless letter strings were selected on the basis of visual similarity to actual words. For this purpose the recognition model was used to predict the most probable, but meaningless letter strings for real words Most of these letter strings were easily pronounceable, as judged by two speakers, 250 had no normal Dutch pronounciation. The latter strings generally consisted of three consonants The first type of meaningless strings will be called 'regular nonwords', the latter type 'irregular nonwords'. Together they totalled 787, making 1500 stimuli presented in blocks of 250.

Presentation

All words and strings were typed in lower-case Courier 10 typeface, while the visual field was 30 x 30 cm with a luminance of $150 cd/m^2$, representative of normal reading situations. The reading distance was 60 cm, at which each letter space subtended a visual angle of 0.25 of a degree. Before each test the subject looked at a blurred fixation spot of a faint grey which did not interfere with the perception of the letters. The location of the spot corresponded to the middle letter of the string. Words and meaningless strings were presented in random order.

Subjects

Thirty subjects participated in the experiment. They responded by pressing one of two buttons corresponding respectively to 'word' and 'nonword'. They all saw each word and string only once. Before sessions started, they practiced on lists of 100 words of four letters. They were instructed to respond as carefully as possible. Trials were completed in two sessions separated by an average of seven days.

RESULTS

How do we measure word knowledge? One way is to list and count the number of words known by a subject. A word can also be considered to be less well known if it takes the subject longer to realize that it is a word. By this reasoning a word is well known according as more subjects recognize it as such and according as they do so quickly and efficiently. On average the subjects recognized 493 of the 713 words, with a standard deviation of 50. The subjects, therefore, seem to form a reasonably homogeneous group. So, on average, a subject knows 70% of all the words, while 37 words, or about 5% were not known to any of the subjects. On the other hand, only 163 words were known by all 30 subjects. Between these extremes are the words known to a varying number of subjects. The familiarity of these words can be defined as the proportion of subjects knowing them. This definition differs from other definitions of familiarity in the literature.

Response times

In Figure 1 response times for correct word identification are plotted as a function of word familiarity. It can be seen, that it takes increasingly longer to recognize a word as such when it is less familiar. The average response time

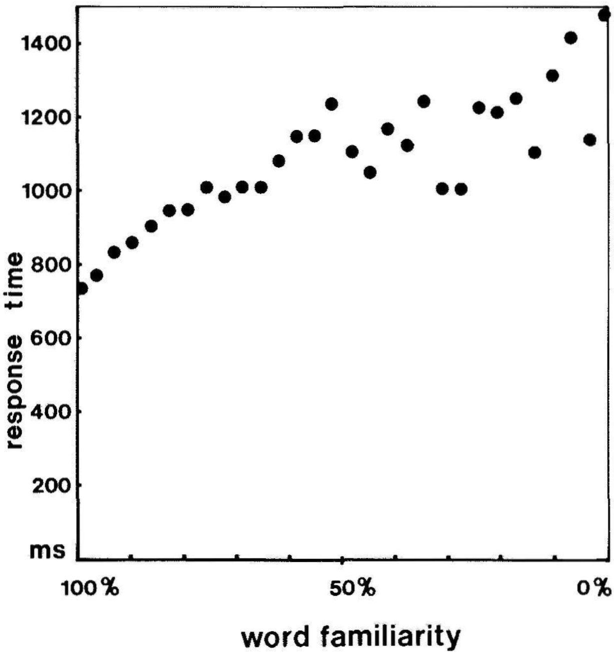


Figure 1. Response times for correct word classifications as a function of word familiarity. Familiarity is the relative frequency of correct classifications by the 30 subjects. Each point represents the average latency for the set of words known by a given number of subjects.

for the words known by all 30 subjects (familiarity 1.0) is 735 ms, that for the least well-known (familiarity 0.033) is 1479 ms, a one hundred percent increase.The standard deviation of these values rises from about 70 ms for the best known words to 200 ms up to a familiarity value of 0.133, after which the standard deviation suddenly increases to about 550 ms, indicating much less homogeneity in this unfamiliar group of words.

Several of the nonwords were classified as words by the subjects. Yet, subjects were more sure about these; there were never more than 16 subjects who classified a nonword as a word.Of the 787 nonwords 542, or slightly less than 70% were were classified correctly by all 30 subjects, who on average had 770 nonwords or 98% correct as compared to the 70% for actual words. Response times for the various familiarity values are shown in Fig. 2.The longest average

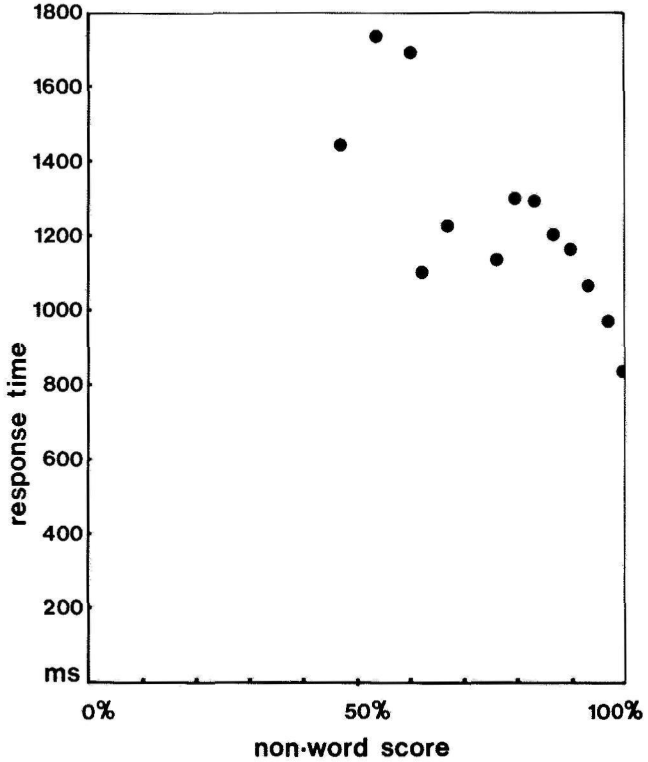


Figure 2. Response times for correctly classified nonwords as a function of the relative frequency with which they were classified by 30 subjects.

time is 1735 ms, decreasing rapidly to 838 ms for the most definite nonwords, with standard deviations of about 150 ms. Consequently, the best known words are classified somewhat more quickly than a definite nonword. If the nonwords are divided into regular and irregular types, a different situation is found. The most definite irregular nonwords are classified in 715 ms, faster, but not appreciably, than the best known words. Classification as a function of word frequency is shown in Fig. 3. Word frequency classes indicated in the figure are based on the 720,000 words counted by Uit den Boogaart (1975) in printed Dutch. Word frequency appears not to influence response time at all, except for the most uncommon words, which are classified 127 ms slower.

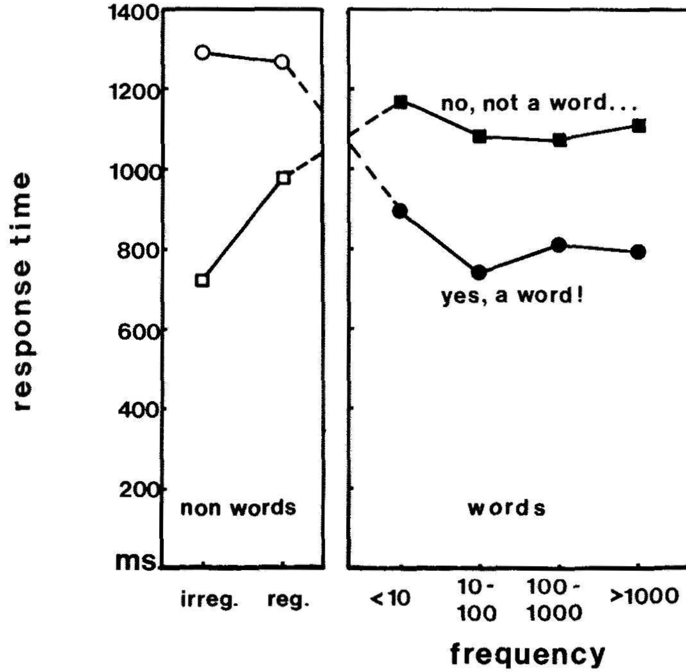


Figure 3. Left: response times for correct and incorrect responses to regular and irregular nonwords.
Right: same for correct and incorrect responses to words as a function of word frequency. Round symbols refer to word responses, square symbols to nonword responses.

These uncommon words include the words which do not appear in counts. Irregular nonwords are classified considerably faster as nonwords than regular ones, which resemble real words, while their identification as words takes longest of all. On average, erroneous responses regularly take 300 ms longer than correct responses, except again for irregular nonwords where the difference is almost twice as large.
In all these analyses geometric means have been calculated instead of arithmetic means. Geometric means are less sensitive to the fairly long response times which occasionally occur, a property which they share with the median (Noordman, 1977; Noordman-Vonk, 1977). Also their associated variances are more stable, but both measures of central tendency preclude the use of additive models.

The differences dicussed are maintained, or even increased if arithmetic means are used.

Accuracy measurements

In the right-hand half of Fig. 4 the percentage of word responses is shown as a function of word frequency. In this case, too, a frequency effect is only discernible for the most uncommon words, namely a drop from 0.95 to 0.62. The left-hand half of Fig. 4 shows the percentage of word responses to nonwords. Of the regular nonwords 2.8% were called words while only 0.6% of the irregular nonwords were so considered. The rank order of accuracy measurements in Fig.4 is almost exactly the reverse of that for the word response times for words and nonwords in Fig.3. If words and nonwords vary in familiarity, it is possible in principle to derive their familiarity distributions. If the familiarity distributions

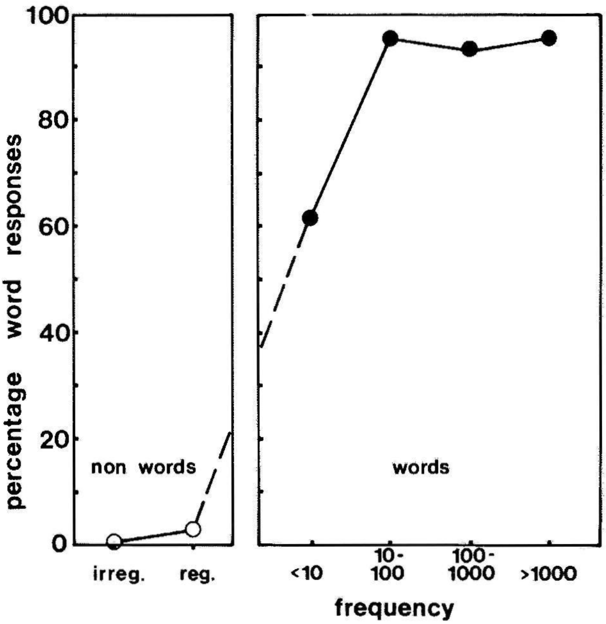


Figure 4. Left: relative frequency of word responses to regular and irregular nonwords. Right: relative frequency of word responses to words as a function of word frequency

overlap, i.e. if some nonwords look more familiar than some words, the subject who bases his response on subjective familiarity, is bound to make some errors. According to the concepts of the theory of signal detectability (Green & Swets, 1966) the underlying distributions can be derived from the probabilities of correct word responses and incorrect nonword responses. Fig. 5 shows these probabilities if both a word and a nonword are called words[1].

If this plot forms a straight line on normal probability axes, the distributions are normal. Since the plot in fact has a linear appearance, a regression procedure was used to estimate the distribution parameters. For the regular nonwords

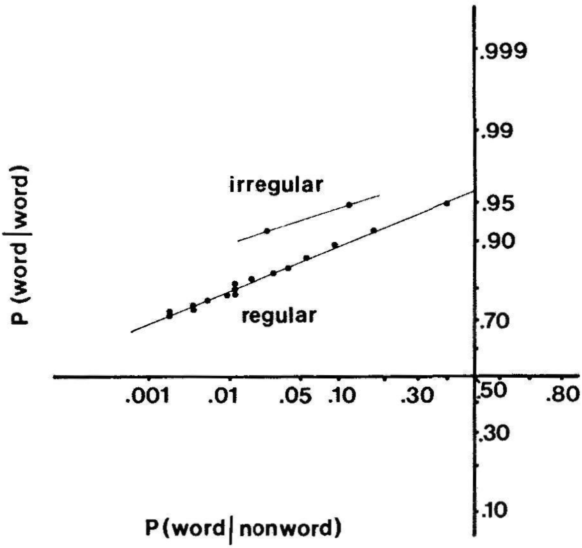


Figure 5. Probabilities of correct word responses plotted against probabilities of word responses to regular and irregular nonwords. The units of both axes are normal deviates.

1) The points in Figure 5 were obtained by taking the cumulative proportions of numbers of words known and the cumulative proportions of numbers of nonwords classified as words by a varying number of subjects. The number of subjects is therefore used as the criterion in signal detectability: strict when all subjects agree, lax when the number of subjects knowing words decreases.

the accuracy of fit of a linear relation, expressed as a proportion of explained variance is 0.991. For irregular nonwords there are only two observations since so few errors were made with them; a linear fit is therefore trivial.
It can be clearly seen that a small shift in the familiarity of words corresponds to a large shift in the familiarity of both types of nonwords. This means that the familiarity of the nonwords must have a much narrower distribution than that of the words. This can clearly be seen in Fig. 6, where the distributions are shown for irregular nonwords, regular nonwords, and words respectively. Words appear to have an enormous spread of familiarity, as compared to nonwords.

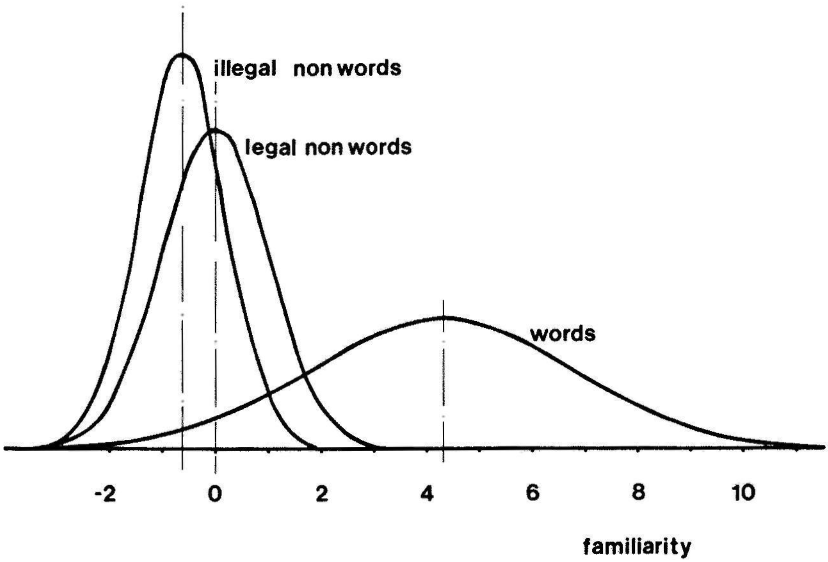


Figure 6. Theoretical familiarity distributions of irregular nonwords, regular nonwords and words, derived from the data points in Figure 5. The unit is the standard deviation of the regular nonwords.

Though there is little confidence concerning the irregular nonwords, whatever there is indicates that their familiarity is still lower than that of regular nonwords, and also less variable.

DISCUSSION

The principal objective of the present study was to assess which three-letter words subjects knew and thus to improve the word vocabulary of the word recognition model developed (Bouwhuis, 1978). In the word recognition model the vocabulary was limited to 541 words of three letters which were all considered to be equally accessible. The present study reveals that subjects effectively know about 493 words and that the words known vary substantially in accessibility. Though the words known vary somewhat from one subject to another, any two subjects knew 426 words in common on average. It can be seen in Fig.1 that if words are less well known, they are responded to more slowly. But these response times stem from subjects who <u>knew</u> the words at the time of presentation. Accessibility, therefore, seems fairly uniform for the subjects; on the whole, subjects know the same words well and they all have difficulties with little known words. It can be concluded from the above findings that the number of words adopted in the large vocabulary was slightly on the high side. Any lower number would reduce the underestimation of the predicted correct responses. However, although the nominal number should be lower, the actual number of words that are accessible might even be greater because not all words are equally accessible. A gradation of accessibility seems to be called for, some words coming quickly, others being difficult to evoke. This situation could be achieved in the model by weighting each word with its familiarity value. The number of possible alternatives increases, allowing greater precision, while the unfamiliar words figure less prominently in the alternatives, thus counterbalancing the underestimation in the predictions. The familiarity values are, therefore, based on experimental data other than those of word recognition as predicted by the model. Since letter perception as defined in the model was also based on separate experiments, the recognition model would need no parameter estimation from word recognition data which it is meant to describe.

Orthographic structure

Pollatsek and Carr (1979) argue that abstract orthographic rules facilitate word perception. They conclude this from a range of 'same-different' experiments in which subjects decided more quickly that two words were the same than that two nonwords were the same. The effect whereby subjects decide more readily that words are identical is called the word superiority effect, which Pollatsek and Carr (1979) associate with orthographic regularity. In our experiment it appeared that the familiarity of words in particular was responsible for variation of word identification response time. This indicates that accessing operations in the reader's word lexicon account for most of the time (Fig. 1). For the present argument it may be assumed that orthographic regularity does not differ systematically for familiar and unfamiliar words, i.e. whether they are well known or not. Then, if orthography was a major factor, unfamiliar words with a normal spelling pattern would have been correctly classified in about the same time as more familiar words. But Fig. 3 shows familiarity to have the largest effect of all for words of the same degree of regularity. However, orthographic regularity does play an important role in the nonword classifications. Irregular nonwords are classified fastest of all types of stimuli and display the fewest errors (left parts of Figs. 3 and 4). Regular nonwords, however, retaining orthographic regularity, are classified more slowly and with more errors than irregular strings. In view of the many facets of orthographic regularity, as exemplified by R. Venezky (1979), orthographic regularity should be related to a process model for word recognition and its role would then probably be greater in the recognition of words in running text.

Word frequency

It is interesting to note that the frequency effect found is restricted to the most unfamiliar words in both response times and errors. Response times vary much more as a function of familiarity, i.e. the extent to which words are known. Frequency and familiarity seem to be distinct properties, the first objective, the latter dependent on the subject's experience, but an experience which seems to be fairly general. Other lexical decision experiments have yielded results showing some word-frequency effects(Rubenstein et al., 1970; Rubenstein et al.,1971; Forster & Chambers, 1973; Stanners, Jastrzembski & Westbrook, 1975). These effects are usually an increase of some 150 ms in the response tires to infrequent, but still well-known words; the number of words presented is small and subjects are usually instructed to respond as soon as possible. The similar frequency effect in the present experiment is 127 ms for words of a much lower frequency than usually employed in the literature, except in the study of Forster and Chambers (1973). The absence of an apparent effect for the higher frequencies may have been caused by the lenient instruction; in difficult perceptual situations the frequency effect is more prone to occur (Richards, 1973).

To summarize, it would appear that the most important factor identified in this experiment on word knowledge is the enormously varying familiarity of words when presented in isolation. This finding means that the original assumption of a limited word vocabulary of equally important words for the recognition model of Bouma and Bouwhuis (1975) is certainly too restricted. The vocabulary would have to be extended and the accessibility of its entries would also need to be modified. That context may facilitate the perception of unfamiliar words seems logical but it implies the existence of very efficient accessing rules.

As previously mentioned elsewhere (Bouwhuis, 1978), the descriptive power of the word recognition model in its present form depends on its application to words of three letters. In longer words the constituent letters cannot be clearly discerned and even length is misperceived. This same restriction is perhaps also applicable to word frequency which may be less effective for three-letter words, of which there are not many. For longer words too, however, much larger effects cannot be expected from word frequency. It would, therefore, appear useful not only to extend the recognition model to longer words, but also to study their familiarity and the effect of their familiarity on recognition.

REFERENCES

Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. Vision Research, 1973, 13, 767-782.

Bouma, H., & Bouwhuis, D. G. Word recognition and letter recognition. IPO Annual Progress Report, 1975, 10, 53-59.

Bouwhuis, D. G. A model for the visual recognition of words of three letters. In: J. Requin (Ed.) Attention and Performance VII, Hillsdale: Lawrence Erlbaum Associates, 1978.

Broadbent, D. E. Word frequency effect and response bias. Psychological Review, 1967, 74, 1-15.

Chambers, S. M., & Forster, K. I. Evidence for lexical access in a simultaneous matching task. Memory and Cognition, 1975, 3, 549-560.

Clarke, F. R. Constant ratio rule for confusion matrices in speech communication. Journal of the Acoustical Society of America, 1957, 29, 715-720.

Forster, K. I., & Chambers, S. M. Lexical access and naming time. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 627-635.

Forster, K. I., & Bednall, E. S. Terminating and exhaustive search in lexical access. Memory and Cognition, 1976, 4, 53-61.

Green, D. M., & Swets, J. A. Signal detection theory and psychophysics, New York: Wiley, 1966.

Kruyskamp, C. Van Dale groot woordenboek der Nederlandse taal, The Hague: Martinus Nijhof, 1970.

Kučera, H., & Francis, W. N. Computational analysis of present-day English, Providence: Brown University Press, 1967.

Linschoten, J. De la Court's frequentietelling van Nederlandse woorden, Report no. 6301, Psychological Laboratory, University of Utrecht, Utrecht, 1963.

Noordman, L. G. M. Inferring from language, Unpublished doctoral dissertation, University of Groningen, Groningen, 1977.

Noordman-Vonk, W. Retrieval from semantic memory, Unpublished doctoral dissertation, University of Groningen, Groningen, 1977.

Pollatsek, A., & Carr, T. Wholistic and rule governed encoding processes in word perception, In:P.A. Kolers, M. Wrolstad and H. Bouma (Eds.) Processing of visible language I, New York: Plenum Publishing Corporation, 1979

Richards, L. G. 'Vexierversuch' revisited: a reexamination of Goldiamond and Hawkins' experiment, American Journal of Psychology, 1973, 86, 707-715.

Rubenstein, H., Garfield, L., & Millikan, J. Homographic entries in the internal lexicon, Journal of Verbal Learning and Verbal Behavior, 1970, 9, 487-494.

Rubenstein, H., Lewis, S., & Rubenstein, M. A. Homographic entries in the internal lexicon: effects of systematicity and relative frequency of meanings, Journal of Verbal Learning and Verbal Behavior, 1971, 10, 57-62.

Rumelhart, D. E., & Siple, P. Process of recognizing tachistoscopically presented words. Psychological Review, 1974, 81, 99-118.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. Frequency and visual quality in a word-nonword classification task, Journal of Verbal Learning and Verbal Behavior, 1975, 14, 259-264.

Thorndike, E. L., & Lorge, I. The teacher's wordbook of 30,000 words, New York: Bureau of Publications, Teacher's College, 1944.

Uit den Boogaart, P. C. Woordfrequenties in geschreven en gesproken Nederlands, Utrecht: Oosthoek/Scheltema en Holkema, 1975.

Venezky, R. L. Orthographic regularities in English words, In: P.A.Kolers, M.Wrolstad and H.Bouma (Eds.) Processing of visible language I, New York: Plenum Publishing Corporation, 1979

4 The Contribution of Letter Structure to the Familiarity of Dutch Three-Letter Words

A reanalysis of the results of a lexical decision study of all 713 Dutch three-letter words revealed that positional letter frequency had a very significant effect. The occurrence of letters in the positions of a word is only partially controlled by orthographic rules, which are more concerned with letter sequences. In addition, positional letter frequency in orthographically regular strings can vary substantially. It appeared also that digram frequency in the three-letter words could be accurately predicted from the occurrence of single letters in corresponding positions and independent combination.

Effects of digram frequency on decision accuracy and decision latency were much smaller than those of single letter frequency.

A mechanism is proposed for lexical access in which word familiarity and letter familiarity act independently.

The lexical organization is described as an associative memory for which the letters of the word function as the address. The working of such a system corresponds to that of the letter confusion model and the logogen system for word recognition. Models of lexical access proposed in the literature usually employ a serial search through lexical memory. It is argued that parallel access can explain the results in a more coherent way.

In a lexical decision study of Dutch three-letter words (cf Chapter 3) the main factor determining the ease with which subjects could distinguish words from nonwords was found to be familiarity. Words were correctly classified as words with probabilities ranging from one to zero and at the same time decision times varied from about 750 ms to 1500 ms. The covariation of these variables gave rise to the concept word familiarity.

Logically word familiarity is determined to a large extent by lexical properties of words; words a subject knows better are more familiar to him. However, the lexical representation of a word is not directly observable, a situation which has led to the study of word frequency as a vicarious measure of word familiarity. Yet for a reader the printed letters of the word are the first agents that give access to the lexical representation, and some part of familiarity might reside in the letter structure of words, which is specific for a given language.

To give an example, very few readers of a Western European language will identify the string *uisge* as a word. When it is printed in its English form, *whisky*, pronounced similarly, very few readers anywhere will not know it. Uisge is a Gaelic word, literally meaning water, and it gives a hint of the typical writing system. The Gaelic writing system possesses many highly idiosyncratic features, probably because it was developed before that of most present-day European languages. So it seems that a good deal of the familiar appearance of printed words is mediated by their letter structure.

In the words involved in the lexical decision study (Bouwhuis, (1979) ; see Chapter 3) several letter structure properties can be analysed owing to the fact that they comprise the full set of 713 Dutch three-letter words and so are a population in the statistical sense of the word. Specifically the frequency of the constituent letters can be determined. In connection with the letter confusion model it is relevant that letter frequency should be position-specific. Three-letter words contain only digrams as higher order letter groups, which makes for a moderate structural complexity. It might be possible in principle that digrams or digram frequency are employed fully in lexical access of three-letter words.

We will first present some details of the letter structure of Dutch three-letter words. Next it will be shown how variables directly derived from the constituent letters affect lexical decision time and accuracy. A provisional and qualitative theory accounting for the results is presented in the discussion. Alternative theoretical proposals will be briefly treated in connection with the results obtained. Although the present findings are not completely unexpected, they provide independent support for the letter confusion model, in particular with respect to the perception of letters in their position.

LETTER AND DIGRAM FREQUENCY

Table 1 presents the frequencies with which the various letters appear in the three positions of Dutch three-letter words. It is seen, that the frequencies vary widely, especially for the middle ltter, which is in 87.4% of all words a vowel. The first letter is a consonant in 84.6% of the words, a terminal consonant appears in 83.3%.
Most three-letter words have consequently a CVC structure. It can also be seen that in all three positions at least two letters do not occur, which reflects elementary properties of spelling in that some letters are not allowed to occur in initial or final position. Letters in medial positions seem to be rather constrained by the neighbouring letters.
At this point it is interesting to speculate on the extent to which the only possible higher order letter group, the digrams depend on the single letter structure presented in Table 1. The simplest hypothesis would be that letters may combine independently to form digrams in words. This leaves no room for spelling rules defined on letter sequences. Formally, the independence hypothesis cannot hold. Complete independence would allow, for example, the string 'bbb' since the words 'bad' (bath), 'abt'(abbot) and 'heb' (have) exist. However, the trend of the data in Table 1 is such that these combinations are made very improbable because of the frequency of the intermediate vowel.
Further the inclusion of a vowel anywhere makes a lot of combinations of three letters pronounceable and lawful spelling forms. In an attempt to check the independence hypothesis we formed all possible pairs with the letters from Table 1 for positions 12, 13 and 23 of Dutch words and calculated their probabilities. These predicted probabilities

Table 1

Letter frequency distributions of initial, middle and final positions of Dutch three-letter words. Data are given in percentages. Words containing the letter combination 'ij' (pronounced as /ɛɪ/) were excluded from the analysis, because it is not clear whether 'ij' may be considered as two letters or one. There are 15 words of the form 'bij' (bee) and 'ijs' (ice). In Uit den Boogaart's list (1975) there are 94 words of the form 'ijle' (thin), 'rijk' (rich) and 'blij' (happy).

letter position

1		2		3	
b	7.3	e	21.0	t	11.1
d	6.6	o	20.1	l	10.1
p	6.5	a	18.5	s	9.8
l	6.5	i	14.3	k	9.7
k	6.3	u	13.5	n	7.2
h	6.3	r	3.5	p	6.3
r	6.2	l	2.7	f	5.6
t	5.5	n	1.4	m	5.5
m	5.5	k	0.6	e	5.5
a	5.0	d	0.6	g	4.6
w	4.3	t	0.4	r	4.2
o	4.2	s	0.4	i	3.6
n	4.1	p	0.4	a	3.2
g	3.8	h	0.4	d	3.1
e	3.8	g	0.4	b	2.5
v	3.2	c	0.4	u	2.4
s	3.2	b	0.4	o	2.0
z	2.8	m	0.3	x	1.5
j	2.8	j	0.3	h	0.8
f	2.5	w	0.1	w	0.7
i	1.3	v	0.1	c	0.4
u	1.1	f	0.1	z	0.1
c	1.1	z	.0	y	.0
q	0.1	y	.0	v	.0
y	.0	x	.0	q	.0
x	.0	q	.0	j	.0

were correlated with the observed probabilities with which the same digrams occurred in Dutch three-letter words. The results are shown in Table 2. These correlations are surprisingly high.

Table 2

Correlations between predicted and observed digram frequency in Dutch three-letter words

	12	.886
digram	13	.772
	23	.910

The lowest correlation is that for the somewhat uncommon digram 13, which is not a proper letter sequence. In this digram the middle letter is lacking, which letter shows the very large frequency difference of vowels versus consonants. This property has a stabilizing effect on the frequencies of digrams 12 and 23 in which the middle letter occurs, but does not affect frequencies of digram 13. Both the other digram correlations are strong indications that letter combinations in three-letter words are independent. It is interesting to note that in a number of cases where observed digram frequency was higher than predicted, some digrams stemmed from foreign intrusions or loan words. For digram 'du-' these were 'duo' (pair) and 'dur' (major key); for the digram 'al-' these were 'ale' (beer), 'alf' (Greek mythological spirit), 'alm' (Alpine meadow), 'alp' (Alp) and 'alt' (alto).

The consequence of the independence property is that it is wasteful for a subject to adjust his perceptual processing to just digrams. There are many more digrams than letters and, in addition, they overlap. Their occurrence can be regularly predicted by a much smaller number of letters in their position.

THE RELATION BETWEEN LETTER FREQUENCY AND DECISION LATENCY

If letters have a varying frequency of occurrence they might be important in mediating word familiarity. This will be checked in two ways. Firstly, familiar letters may decrease the word-nonword decision time; secondly, they may increase the accuracy of word-nonword decisions. Such analyses have been carried out for the three letter positions separately. Table 3 shows the correlations between letter frequency and decision latency for words and regular and irregular non-words. A definition of the stimulus types can be found in Chapter 3. In the analysis arithmetic means were used in preference to geometric means employed in Chapter 3. The main reason is that, unlike geometric means, arithmetic means allow linear combinations. As noted in Chapter 3, the results are substantially the same apart from a slight shift in absolute level. Table 3 shows the correlations between positional letter frequency and response latencies[1].

Table 3

Correlations between positional letter frequencies and lexical decision latencies. All correlations are significantly different from zero beyond the p = .002 level, except where indicated otherwise.

stimulus type	letter	correct response		incorrect response	
WORDS	1	-.633		.569	
	2	-.589		-.076	*N.S.*
	3	-.668		.549	
REGULAR NONWORDS	1	.642		.096	*N.S.*
	2	.466	$p < .05$	.203	*N.S.*
	3	.709		-.035	*N.S.*
IRREGULAR NONWORDS	1	-.202	*N.S.*	-.160	*N.S.*
	2	.718		.049	*N.S.*
	3	-.406	*N.S.*	.288	*N.S.*

[1] Note that the middle letter was not less discriminable than the initial and final letter since presentation in the lexical decision task was foveal.

The frequency of single letters can account for up to 40% of the variance in lexical decision latency for the real words; when letters are frequent the decision latencies decrease. Interestingly, the frequency of the constituent letters also delays erroneous decisions, as appears from the positive correlations with error latencies. This implies that when a word is not identified as a word, the error response is delayed more when the word has frequent letters than when it has infrequent letters. The reverse relation holds for regular nonwords, except that error latencies are unrelated to letter frequency. When a pronounceable nonword comprises frequent letters it is only slowly rejected as a word.

Irregular nonwords show hardly any effects of letter frequency. Their structure should by itself be sufficient for rejection without further checks on lexical representation. Yet the frequency of the middle letter correlates positively and highly with rejection latency. As remarked before, three-letter strings can usually be made pronounceable by insertion of a vowel. When the vowel is absent, especially in the middle positions, the string will appear to be irregular and unpronounceable. In fact, most irregular nonwords consist of three consonants, indicating that a low frequency middle letter, a consonant, will produce a fast decision. The effect of letter frequency on latency can also be shown in a more detailed picture, revealing the relationship with word frequency. To this end, letters were divided into a high frequency and a low frequency group. For all three positions a suitable criterion for classification was 1/26 or 3.85% which would be the value of equal frequency. For each of the three positions this yielded relatively well-spread frequency classes. The effects of high and low frequencies of initial, middle and final letters are shown in Figures 1-3, which are adapted from Fig. 3 in Chapter 3. These latency differences are shown as a function of word frequency and nonword regularity. In the figures all data points referring to less than 50 latency measurements have

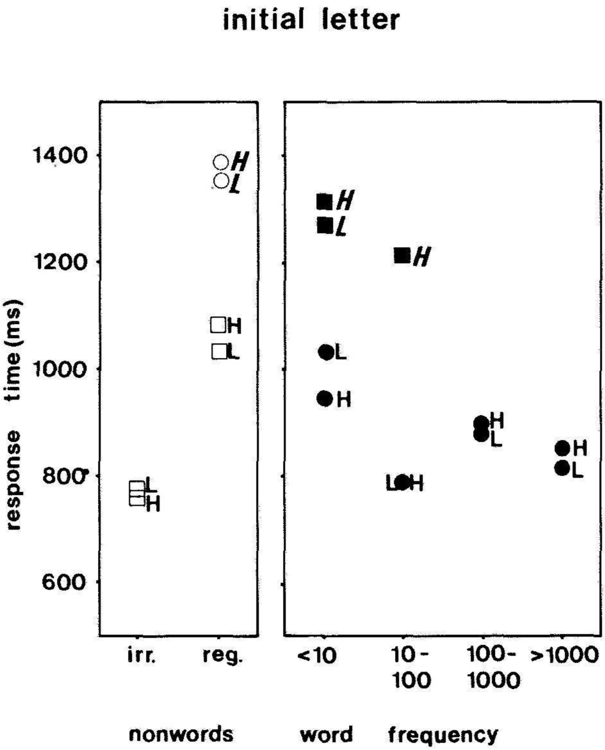


Figure 1

Decision latencies as a function of word frequency (right) and type of nonwords (left). Frequency of the initial letter is the parameter, (H: High; L: Low). Round symbols denote word responses, square symbols nonword responses; incorrect responses in italics. All latency differences (H-L) shown are significant ($p < .05$) except for correct responses to words with a frequency between 10 and 100.

been omitted. Since the latency distributions were very skewed and had large variance differences the requirements for an analysis of variance could not be met. Therefore latency differences were analysed with t-tests for one word frequency group or nonword type for correct and incorrect response separately. In these cases variances were comparable, but it should be realized that t-tests requirements, as in

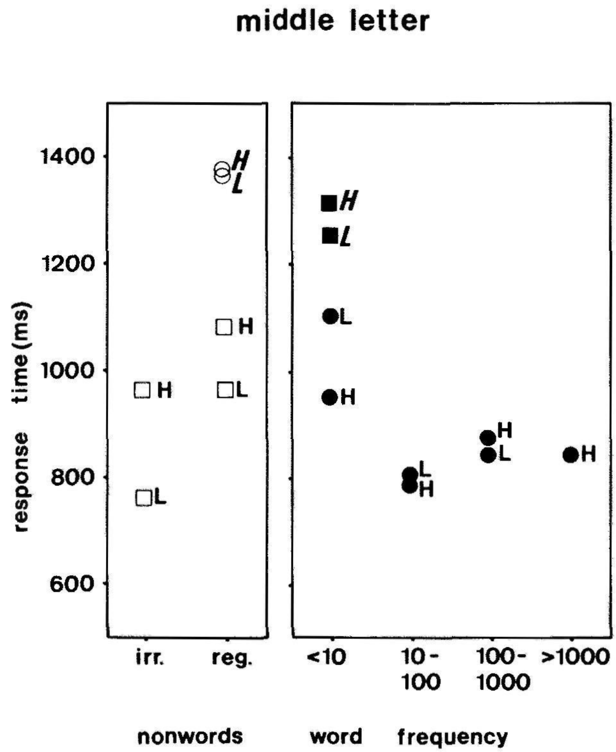


Figure 2

Same as figure 1 for the frequency of the middle letter. All latency differences (H-L) are significant (p < .05) except for word responses to regular nonwords. There was only one word with a frequency higher than 1000 having a low frequency middle letter.

all other lexical decision experiments, were also not met. All differences mentioned in the captions are significant to p less than .05; the location of the significant differences is reported in the figure captions. The effects of differential letter occurrence in words with a frequency exceeding 10 (on a total of 720.000 words counted by Uit den Boogaart, 1975) were unsystematic over all letter positions, and mostly small. In all cases these date relate

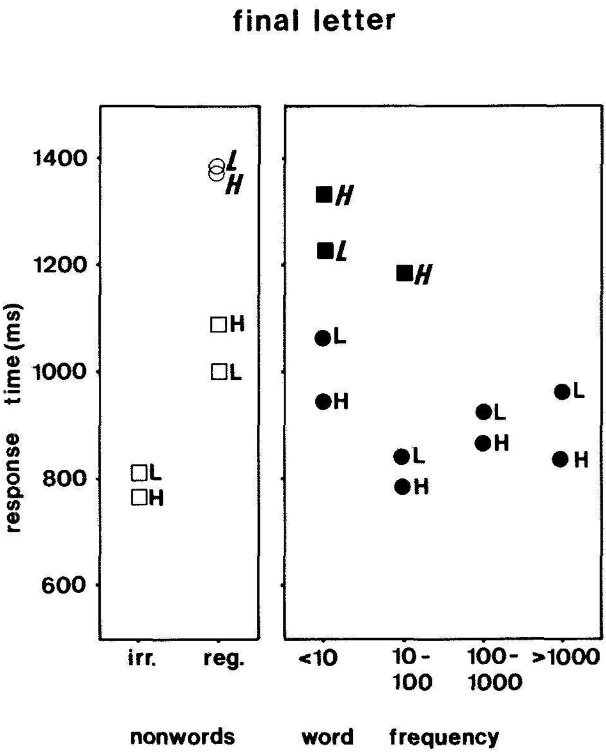


Figure 3

Same as figure 1 for the frequency of the final letter.
All latencies (H-L) are significant except for word responses to regular nonwords.

to only few words, ranging from 1 to 25. Effects of letter frequency in these words could be dependent on the particular sample of words. However, the results are completely clear-cut for the low-frequency words and the regular nonwords. For the rare words latency differences between words with frequent letters and word with infrequent letters are significant for correct decisions and errors in the expected direction. For regular nonwords there is no relation with letter frequency in errors, but a nonword response takes longer when their letters are frequent. This pattern holds for all three letter positions. As expected, there is a large latency decrease of correct rejection of irregular nonwords (Fig. 2) when the middle letter is a low frequency consonant.

THE RELATION BETWEEN LETTER FREQUENCY AND DECISION ACCURACY

As has been done in the previous study (Chapter 3) the same analysis can be repeated for accuracy measurements. Table 4 shows the correlations of letter frequency with the probability of a word response for the several stimulus types. For the nonwords this is actually a measure of inaccuracy. The correlations are all significantly different from zero at $p < .01$ level, except where otherwise indicated.

Table 4

Correlations of letter frequency with probability of word responses

letter	words	regular nonwords	irregular nonwords	
1	.624	.703	.075	*N.S.*
2	.438	.646	.332	$p = .05$
3	.687	.759	.402	

The positive correlations indicate the greater tendency to respond 'word' when the letter string is composed of frequent letters. As for latencies, this tendency is smallest for irregular nonwords. For the words and the regular nonwords the correlations are of the same level, but also show a similar trend for the letter positions as in the case of latencies. This is true in spite of the fact that the instruction stressed accuracy rather than speed, by which larger effects would be expected in latencies than in accuracy. This might indicate a strong coupling between the two measures; which was also observed in a number of word recognition tasks (Bouwhuis, Schiepers, Schröder & Timmers,1978).

When, as has been done for latencies, the letters are divided into frequent and infrequent groups, their effect on word score can be analysed for the various stimulus types. The results are in line with the previous ones. For the most frequent words the differences in word score are not significant for letter frequency. The test used was the binomial test, which is not quite applicable because of differences between words and between subjects. For most types of stimuli these differences are not large, except for the most infrequent words where subject differences are most prominent. Fortunately, these differences make the test only more conservative (Bouwhuis, 1973).
In the frequent words, (f>10), letter frequency has a very small and insignificant effect on the probability of word responses. It appears that for the rarest words, subjects give significantly more word responses ($2.5 < z < 10$) when words have frequent rather than infrequent letters. The same is true of the regular nonwords, though the effect is much smaller. The differences are consistently of the same size as the standard error. Though strictly speaking the differences are not significant, an effect of letter frequency is indicated here as well. The irregular nonwords show no systematic differences, but error were very rare.

THE EFFECTS OF DIGRAM FREQUENCY

If, as has been shown, digram frequency can be accurately predicted from single letter frequency, the effects of digram frequency should be predictable to some extent if digrams are processed as combinations of single letters. Independent combinations of letters weaken appreciably the frequency variations of the resulting letter pair on average. Thus, while several letters occur in more than 100 words, the most frequent digram occurs in only 18 words. As a consequence, correlations of digram frequency with latency and with accuracy would be expected to be lower, but in the same direction as those with single letter frequency. A summary is given in Table 5.

Table 5

Correlations of digram frequency in Dutch three-letter words

	WITH DECISION LATENCY		WITH WORD SCORE
digram	correct word responses	incorrect responses	words
12	-.271	.349	.303
13	-.102	.291	.185
23	-.344	.240	.342

The much smaller size of the correlations for digram frequency make it a much less interesting measure than frequency of single letters in their position. Digram analysis for nonwords is impeded by the fact that most nonwords contain digrams which do not occur in Dutch three-letter words and consequently cannot yield a digram frequency effect. Apparently, the correlations are low despite the fact that digrams represent a much larger part of a three-letter word than single letters. It is also surprising that the letter frequency effect occurs when the subject has sufficient time for his response to base it on higher order features.

DISCUSSION

The results of the present analyses throw light not so much on the role of one particular factor mediating word familiarity as on the interplay between visual or graphemic factors and lexical representation. In Figures 1-3 it is shown that for words occurring more than ten times in Uit den Boogaart's count (1975) decision latencies are not systematically affected by letter frequency. These words are faster classified than rarer words or regular nonwords. It is therefore probable that relatively frequent words permit a rather fast and efficient access to their lexical entry.

Familiarity would seem to be mediated by the frequency of those words, by which accessing procedures are modified or representations more extended. In signal detection terms this kind of familiarity might be interpreted as a lexical criterion shift permitting less information to evoke the word. The rare words, however, many of which are much less well known, still carry properties of Dutch spelling regularities as expressed in letter content. These letter properties can compensate partly for the lower accessibility of the corresponding lexical entries. It appears from the results that words are correctly classified faster when they are comprised of frequent letters than when they have infrequent letters. Conversely, when rare words are erroneously classified as nonwords, frequent letters delay this decision systematically. Consequently, in uncertain situations, which frequently arise in lexical decision tasks, positional letter frequency makes up for a good deal of word familiarity. Regular nonwords, in themselves sources of uncertainty, are more quickly rejected as words when their letters are infrequent. It might be expected that a word decision for a regular nonword might correspondingly be speeded up by frequent letters. This is certainly not apparent from the data, which might indicate that letter frequency is irrelevant for those decisions. Yet letter frequency might affect decision latencies for nonwords in two ways. First, of course, frequent letters predispose the subject to give word responses and this should result in a latency decrease. However, when a meaning cannot be found, the familiar letter might still give the impression that the string is a word. This tendency might lead to a word response which is delayed by the vain search for meaning. Evidence of such an effect is provided by frequent remarks of subjects that they believed they had encountered the string -or, as they called it: the word,- some time ago in a crossword puzzle. These familiar nonwords might aptly be called crosswords. The two effects might compensate each other, causing the same response latencies for frequent letter nonwords and rare letter nonwords.

It is almost impossible, in view of the large variance of these long response times, to distinguish these two effects in the data, or to establish the size of them.

A processing mechanism for lexical access

It is proposed here that by virtue of the nature of lexical access constituent letters of a word are used as an access structure to retrieve the lexical entry and the connected meaning or semantic attributes. If the constituent letters were indeed the sole access structure, the system would be errorless in that no false alarms would be produced: the letters of a word should not give rise to another word with a different meaning. Naturally, this presupposes perfect stimulus conditions, as are present in a lexical decision study. An 'error' which might occur is that an address fails to evoke the lexical representation, for example when the occurrence of the word is improbable in the context, or the word meaning is forgotten, or cannot be retrieved. Both phenomena, the absence of false alarms and forgetting, seem realistic. Such an access mechanism does formally correspond to an associative memory. In an associative memory entries are searched in parallel. The entries consist of a file of features and basically the address to locate an entry is part of the feature list itself. The most elementary operation is therefore a check on whether the entry exists or not. The associative memory is sometimes called content-addressable, because the entry can be found by its content, e.g. the letters of a word. The interesting feature of an associative memory is the simultaneous search on all of its entries, which makes a short search duration possible.

This mechanism may be contrasted with models proposed in the literature. Lexical search was assumed to be sequentially organized among others by Landauer and Freedman (1968) who found that decisions on membership of a smaller set (dog = animal?) were faster than decisions on membership of a larger set (dog = living thing?). Parallel search in larger sets would also take more time if search times per entry were stochastic variables. If the search is sequential and

self-terminating the whole set has to be searched before it can be decided, that an item does not belong to it. Temporally a parallel search would again have the same properties. Sequential search for lexical decisions has been advocated by Rubenstein, Garfield and Millikan (1970); Rubenstein, Lewis and Rubenstein (1971a); Jastrzembski and Stanners (1975) and by Forster and Bednall (1976). The implication of the sequential model is that entries are searched at tremendous rates of thousands of words per 100 ms, which is a representative experimental effect. This idea is qualified by most authors by stating that only subsets are searched, which may be defined by graphemic or phonemic codes of the perceived word. The idea of a content-addressable memory is compatible with the logogen system of Morton (1969) or the word recognition model of Estes (1977). Logogens, or memory vectors of words, are organized in parallel and receive each item of information of the stimulus pattern which refers to it.
When the pattern does not give rise to errors, only the correct logogen will receive information and release the word response. This scheme seems appreciably less wasteful than serial search. Parallel search models have been proposed by Novik (1971) and by Snodgrass and Jarvella (1972), though not elaborated in much detail.

Phonological recoding

All words, rare and frequent ones and regular nonwords, have a lawful spelling and are easily pronounceable.
Phonological coding might be operative in both rare and frequent words but cannot explain differences between them, though phonological coding might be closely related to positional letter frequency.
In the literature arguments in favour of phonological coding are usually derived from performance on nonwords (Rubenstein, Lewis & Rubenstein, 1971b). For the present analysis the arguments in support of graphemic letter factors are wholly derived from properties of real words. This difference suggests that task requirements may induce types of coding

which may be useful in some tasks but inefficient in others. Phonological recoding seems especially useful in the so-called double lexical decision tasks in which the subject has to decide whether a pair of letter strings are both words or both nonwords. Types of stimulus combination and context or other stimuli can indeed influence decision speed but seem to reveal the limits of coding capacities rather than their functioning in less demanding tasks (Davelaar, Coltheart, Besner & Jonasson, 1978).

The contribution of single letters and digram

In a word recognition study McClelland and Johnston (1977) found that digram frequency had no effect at all on several report measures, but single letter positional frequency had an effect as large as that of word frequency. This finding was preceded by a related result of a study by McClelland (1976) who presented words in mixed case.
As expected, a word like 'fare' was perceived more accurately than 'fArE'. The observed phenomenon that the mixed case 'fArE' was more accurately perceived than the nonword 'lare' (in same case) essentially excludes the operation of higher order units or digrams because of their complete unfamiliarity in the mixed case.
There are two lexical decision studies in which letter frequency has been studied. Stanners, Forbach and Headley (1971) presented words and nonwords of a CVC structure and found that words were classified correctly 80 ms faster when the initial and terminal consonants were frequent rather than infrequent. In the same conditions nonwords were classified slower by the same amount. As was found in the present study, no effect of consonant frequency was obtained in CCC strings. A replication of the experiment with CCVCC words and nonwords (Stanners & Forbach, 1973) yielded the same effect for words, while that for CCVCC nonwords was increased to 213 ms. In both experiments errors were too rare for a meaningful analysis. Likewise CCCCC nonwords were classified 60 ms slower when their initial and terminal consonants were frequent letters. Since the latter differed from words only

by the middle consonant they must have looked more familiar than CCC nonwords. Stanners et al. (1971, 1973) argue that the initial and terminal consonants are coded, possibly phonologically, and in this way determine a subset of possible lexical entries to be searched according to the serial comparison scheme (Rubenstein, Lewis & Rubenstein, 1971b). It should be realized, however, that initial and final letters are as a rule visually relatively prominent. Evidence supporting the informative value of the first letters was found by Broerse and Zwaan (1966). They maintain that beginnings of words are more informative, and that retrieval occurs by sequential letter patterns, the beginning of the word being the obvious starting point.
A range of word recognition experiments in which digram frequency in words was varied (Biedermann, 1966; Broadbent & Gregory, 1968; Broadbent & Gregory, 1971) did not reveal large or svstematic effects and these, in turn, might be due to the incidental structure of low frequency English words.

These experiments also led to the conclusion that detection of a letter does not influence detection of other letters. Finally, the results reported here are supported by findings of Mason (1975, 1978) and Mason and Katz (1976). Mason defines spatial redundancy as the position-specific letter frequency in words of a given length. Some features of spatial redundancy are determined by orthography, the letter 'x' cannot be an initial consonant except for Greek derivatives, and 'q' and 'v' cannot be in the terminal position. Letter sequences as such are not determined by spatial redundancy but are controlled by orthographic rules. An interesting finding on spatial redundancy (Mason, 1975) was that poor readers did not employ it, while good readers did in tasks where a target letter had to be found in a single string or a number of strings varying in spatial redundancy. There might also be letter thresholds to the effect that for certain letter positions less information is needed for frequently occurring letters. This is analogous to the word threshold in the logogen system (Morton, 1969). Such a

mechanism would not be unduly complicated as regards the number of letters and positions contrasted with the number of words a reader knows.
Numbers and positions of digrams would require a much more detailed and complicated recognition system. At present our letter confusion model does not account for position-specific frequency effects, but inclusion would not invalidate the basic assumptions. Essentially its basic assumptions of position-specific letter processing are completely in line with possible effects of spatial redundancy, but do not leave room for specific digrams as elementary units of word recognition. It might still be asked whether there are letter combinations which may behave like a unit. For the present experiment it was noted that words containing the combination 'ij' were omitted. In Dutch this letter combination is written as one letter, resembling the y, but it is usually printed as two letters.
It might be possible that it is conceived as a unit together with a limited number of frequent digrams. The 'ij' is a probable candidate, also because of the similarity between the two, relatively slender, letters.

Finally an important condition for a letter combination to become a unit in perception is its independence of position. It should be possible to occur anywhere within the word retaining its properties, which seems to hold for the Dutch 'ij'. Also the English 'th' may be perceived as a unit, but it is still written as two letters.
A problem with the unitary form of letter combinations is still segmentation, which should be strongly impeded by the unit. For example, what about 'hothead', where 'th' straddles the combination of 'hot' and 'head', just as in 'pothole'?

Letter frequency has been taken to be position specific and was shown to be operative in three-letter words for Dutch and for English, but also for initial and final consonants of five-letter words (Stanners & Forbach, 1973).
But what about word combinations like armchair, seaweed or windmill? It is improbable that different letter frequency

in the middle positions of longer words has large effects. The interesting observation to make is that in the case of armchair, for example, the initial letters could access the lexical entry for arm, and the final letters that of chair, which would be easy for a (letter) content-addressable system. The middle letters would then be less helpful because of the segmentation problem. This brings us really to a semantic problem, regarding the existence of a lexical representation for 'armchair'.

Do readers really see the word arm in armchair?

Some evidence on this point comes from a study reported by Murrell and Morton (1974) who found in longer words that subjects actually did perceive morphemes. In an ingenious series of experiments by Taft en Forster (1976) the general finding was that compound nonwords whose first constituent is a word (FOOTMILGE)take longer to classify as a nonword than nonwords whose first constituent is not a word (TROWBREAK). This result indicates that a word first syllable of a nonword may evoke the lexical representation of that word, delaying the nonword response. But if the last syllable of a nonword is a word,the decision time is unaffected.

Taft and Forster (1976) conclude that readers indeed segment polysyllabic words and that lexical access starts with the first syllable. Interestingly, though employing a search type explanation they note that a content-addressable memory may also account for the results. Methodological weaknesses in the experiments make it doubtful, however, that lexical access invariably takes place via the first syllable. The most important point is the control of eye movements of the subjects. Polysyllabic strings were usually eight to nine letters in length and the presence of a fixation point is not mentioned. Subjects could, therefore, fixate first the initial part of the string and then shift their eyes to the last part in the 500 ms. presentation time. This would always favour the first part.

Summarizing, however, also these results suggest that parts of words can be directly matched with lexical representation.

It would therefore seem interesting to investigate the generalizability of the role of single letters and their frequency in longer words and word combinations. For initial and final parts the model of lexical access which is proposed here might still provide a valid description.

REFERENCE NOTE

Bouwhuis, D. G. Het klemtoonoordeel, persoonlijk of eenstemmig? Report no. 253, Institute for Perception Research, Eindhoven, 1973.

REFERENCES

Biedermann, G. B. The recognition of tachistoscopically presented five-letter words as a function of digram frequency. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 208-209.

Bouwhuis, D. G. Word knowledge and letter recognition as determinants of word perception. In: P. A. Kolers, M. Wrolstad and H. Bouma (Eds.) The processing of visible language I, New York: Plenum Publishing Corporation, 1979

Bouwhuis, D. G., Schiepers, C. W. J., Schröder, U. O., & Timmers, H. The temporal structure of word recognition responses. IPO Annual Progress Report, 1979, 13, 83-93.

Broadbent, D. E., & Gregory, M. Visual perception of words differing in letter digram frequency. Journal of Verbal Learning and Verbal Behavior, 1968, 17, 569-571.

Broadbent, D. E., & Gregory, M. Effects on tachistoscopic perception from independent variation of word probability and letter probability. Acta Psychologica, 1971, 35, 1-14.

Broerse, A. C., & Zwaan, E. J. The information value of initial letters in the identification of words. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 441-446.

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. T. Phonological recoding and lexical access. Memory and Cognition, 1978, 6, 391-402.

Estes, W. K. On the interaction of perception and memory in reading. In: D. Laberge and S. Jay Samuels (Eds.). Basic processes in reading. Hillsdale, Erlbaum Ass., 1977.

Forster, K., & Bednall, E. S. Terminating and exhaustive search in lexical access. Memory and Cognition, 1976, 4, 53-61.

Jastrzembski, J. E., & Stanners, R. F. Multiple word meanings and lexical search speed. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 534-537.

Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: category size and recognition time. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 291-295.

Mason, M. Reading ability and letter search time: effects of orthographic structure defined by single letter positional frequency. Journal of Experimental Psychology, General, 1975, 104, 146-166.

Mason, M., & Katz, L. Visual processing of nonlinguistic strings: redundancy effects and reading ability. Journal of Experimental Psychology, General, 1976, 105, 338-348.

Mason, M. The role of spatial redundancy in grapheme recognition: perception or inference? Journal of Experimental Psychology, Human Perception and Performance, 1978, 4 662-673.

McClelland, J. L. Preliminary letter identification in the perception of words and nonwords. Journal of Experimental Psychology, Human Perception and Performance, 1976, 2, 80-91.

McClelland, J. L., & Johnston, J. J. The role of familiar units in perception of words and nonwords. Perception and Psychophysics, 1977, 22, 249-261.

Morton, J. The interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

Murrell, G. A., & Morton, J. Word recognition and morphemic structure. Journal of Experimental Psychology, 1974, 102, 963-968.

Novik, N. Parallel processing in a word-nonword classification task. Journal of Experimental Psychology, 1974, 6, 1015-1020.

Rubenstein, H., Garfield, L., & Millikan, J. Homographic entries in the internal lexicon. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 487-494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. Homographic entries in the internal lexicon: effects of systematicity and relative frequency of meaning. Journal of Verbal Learning and Verbal Behavior, 1971a, 10, 57-62.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. Evidence for phonemic recoding in visual word recognition. Journal of Verbal Learning and Verbal Behavior, 1971b,10, 645-657.

Snodgrass, J. G., & Jarvella, R. J. Some linguistic determinants of word classification times. Psychonomic Science, 1972, 27, 220-222.

Stanners, R. F., Forbach, G. B., & Headley, D. B. Decision and search processes in word-nonword classification. Journal of Experimental Psychology, 1971, 90, 45-50.

Stanners, R. F., & Forbach, G. B. Analysis of letter strings in word recognition. Journal of Experimental Psychology, 1973, 98, 31-35.

Taft, M., & Forster, K. I. Lexical storage and retrieval of polymorphemic and polysyllabic words. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 607-620.

Uit den Boogaart, P. C. Woordfrequenties in geschreven en gesproken Nederlands, Utrecht: Oosthoek/Scheltema & Holkema, 1975.

5 A Retest of the Word Recognition Model with a Varying Knowledge Vocabulary

A lexical decision study of all Dutch three-letter words revealed that they varied substantially in the degree to which they were known. This result motivated a modification of the vocabulary which the letter confusion model employed. First, the vocabulary was extended to all existing words of three letters, and second, all words were assigned a familiarity weighting, reflecting the degree to which they are known.
Despite the considerable extension of the vocabulary, which, without familiarity weigthing would lead to increased underestimation of correct scores, predictions were slightly better for correct scores and correct letters in responses. Deviations that remained can be ascribed to subject differences in word knowledge and letter recognition.
Incorporation of the word knowledge effect in the model is analogous to that in the logogen model and the multicomponent model and thus may be interpreted as a decreased criterion for better known words.

The view that familiarity of words assists in their recognition is as old as reading research itself. The classical reference is that of Cattell (Boring,1950) who found around 1885 that more letters could be perceived in a short time when they formed a familiar word. Similarly, he showed that reading time for languages varied as a function of familiarity of the language, a phenomenon of which the readers themselves were unaware. As was made plausible in the Introduction, (Chapter 1) familiarity with the responses was an important factor in the recognition experiments of Erdmann and Dodge (1898), the results of which led to the whole word theory of word word perception. As such, familiarity is a rather vague term and it was in the New Look theories of the fifties that endeavours were made to give it a more circumscribed meaning. Sources of familiarity were variously attributed to the value system of the reader, the frequency of occurrence of the stimulus, or its expectancy in a given context. The last two factors essentially complement each other in the sense that frequent words are the most likely to occur.

The distinction is preserved in the terms stimulus set and response set. In a stimulus set the sensory system is thought to be sensitized, as a consequence of frequent exposure to common words. In a response set it is assumed that the subject has the tendency to respond with more frequent words, more or less independent of stimulus factors, but corresponding to context factors. The mechanisms by which such facilitation might occur have been reviewed by Broadbent (1967) who concluded that readers accept less information for common words. In this way the better performance for more frequent words is attributed to a criterion shift. Current models of word recognition account for the word frequency effect in that way, like the logogen model of Morton (1969) and the multicomponent model of Rumelhart and Siple (1974). This aspect will be treated in more detail in the discussion of this chapter and in Chapter 6, in which both models are reviewed in quantitative detail.

In contrast, the present version of the letter confusion model does not take frequency effects into account, for which several reasons were mentioned earlier(Chapter 2; Bouwhuis & Bouma,1979). Basically the arguments relate to the insufficiency of word frequency to explain the constitutive process of word recognition and because of the uncertainty concerning the extent to which it reflects subjective familiarity of words. The letter confusion model contains the assumption that a reader knows the words of his language and that by this knowledge words are familiar units, as distinguished from other letter strings. Such a view implies an all-or-none representation of word knowledge; a reader knows a word or does not. Implementing a word recognition model with such a vocabulary can have either of two consequences.

If the word vocabulary is limited, some letter strings will erroneously not be conceived as words by the model. Correspondingly, since the number of words is small, superior performance will be observed for words contained in the vocabulary, which may not be realistic.

If the vocabulary is large, larger than the reader's vocabulary, accuracy will be increased in that predictions can be generated for all possibly occurring words. Predicted recognition performance, however, will drop as a result of predicted response competition. In both other word recognition models, as presented in the literature, such a situation cannot easily arise. In the logogen model parameters of context and frequency are derived from the observed word recognition and word confusion probabilities themselves and not established beforehand. In addition, the model does not specify individual responses to individual stimulus words. This same reasoning applies to the model parameters of the multicomponent model; estimates are derived from the experimental results by which deficiencies in the data description can be compensated for by adjustment of the parameters. In the letter confusion model neither recognition nor knowledge parameters are estimated from the word recognition results, but are derived from letter recognition experiments and word counts respectively. Consequently, if the vocabulary is too large or too small the model fit cannot be improved by adjusting the letter recognition parameters.

In fact it was found that the predictions of correct report probability were too low (Chapter 2). It appeared,too, that a larger underestimation occurred when the vocabulary was enlarged to account for more responses. An improvement of the model should retain the representativeness of the vocabulary, but also counter the concomitant underestimation of the correct scores. Such a mechanism was proposed in the discussion of a lexical decision study (Chapter 3). This study involved all Dutch three-letter words which could be found in a representative dictionary. Words appeared to differ substantially in the degree to which they were known to the subjects. Knowing a word is operationally defined here as the (probabilistic) occurrence that the subject decides within the observed response time that the presented string, forming a word in the dictionary, is indeed a word.However, if a subject decides that the string is not a word, he may, given sufficient time, realize that it is a word after all.[1] These occurrences have been reported occasionally by subjects. This observation permits the inference that words vary in familiarity, or in the ease with which they can be accessed. Thus the lexical decision study yields direct estimates of how large the probability is that an isolated word will be identified as a word from a single presentation under perfect stimulus conditions.

THE RELATION BETWEEN LEXICAL DECISION AND WORD RECOGNITION

The main question is now to what extent the results of lexical decision studies can be employed in a recognition model. In lexical decision a clearly seen word must be checked on lexical representation and the subject may be careful to ensure that the presented string indeed has a meaning. He may, then, be aware of the meaning, but he may also realize that he has forgotten it.
Also the presence of so-called pseudowords, nonwords resembling real words in a lexical decision experiment may induce the subject to a 'deep' lexical analysis.

In a word recognition study imperfectly seen letters may fit the letter representation of several lexical entries, but may also be conceived as meaningless strings. In the model it is supposed that the latter are rejected and the criterion for rejection might be the degree of familiarity, part of which is being mediated by the constituent letters (Chapter 4). For word recognition it is sufficient to realize that the perceived letter string is a word - or forms part of the vocabulary, while an analysis of meaning is not required. This has been mentioned as the basic operation in a content-addressable, or associative memory (Chapter 4). So lexical decision seems, therefore, to require a more extensive analysis of the lexical entry than does word recognition. This may be reflected in the response times obtained in lexical decision experiments, which are usually much longer than for word recognition (Bouwhuis,

1) This was scored as a 'no' response, corresponding to the observed latency

Schiepers, Schroder & Timmers, 1978). The common factor relating to both tasks may be termed access to lexical entries. Accessibility of different lexical entries is variable, as inferred from latency data and accuracy data, and may be expected to be a rather permanent property of words. A word which has a low accessibility requires more time for the reader to decide that it is a word and will not always be identified as a word. But such a word would neither be probable as a response in a word recognition task.

A RECOGNITION MODEL WITH A VARYING KNOWLEDGE VOCABULARY

The correspondence of word accessibility in both tasks will be taken to form the basis for a word vocabulary in which words are not equally well known. The simplest assumption is that a word can be identified as part of the vocabulary with the same probability that it is decided to be a word in a lexical decision task. This representation of word knowledge can immediately be incorporated in the model by assigning to each word its familiarity value obtained in the lexical decision experiment (cf. Chapter 3) which is the probability of a correct word decision.

The recognition model generates a probability that a particular combination of three letters will be seen on a trial. In the model with a variable knowledge vocabulary this probability is weighted with the familiarity value, which is a number between 1 and 0 for words and 0 for nonwords (cf. Chapter 3 and table 1). Nonwords are predicted by the model not to occur, a fact which is almost true. Further, word responses to nonwords in lexical decision tasks take almost twice as much time as word responses to words, by which this tendency would not be expected to become important in the relatively short response time of word recognition. For the prediction of the final response the weighted probabilities are normalized (constant ratio rule) in order to add up to 1. This results in a predicted probability for each word response. The operation of word knowledge in the revised

Table 1

Comparison of the operations of the two vocabularies in the letter confusion model for the word 'eed'. Only those nonwords are shown whose probability exceeds .05. Usually there are far more response alternatives. It can be seen that the effect of familiarity on number of predicted responses is restricted for low probability responses.

stimulus word: eed (oatn) at $+1.75^{\circ}$

541 WORD ALL-OR-NONE VOCABULARY

perceived letter string		probability of $(l_1 l_2 l_3)$	word selection	weighted probability		response probability (×1/.472)	observed probability
cad	(-)	.079	0	.0		-	-
ead	(-)	.087	0	.0		-	-
eed	(oath)	.450	1	.450		.953	1.00
end	(end)	.019	1	.019		.040	-
oud	(old)	.001	1	.001	.472	.002	-
vod	(rag)	.002	1	.002		.004	-

713 WORD VARYING KNOWLEDGE VOCABULARY

perceived letter string	probability of $(l_1 l_2 l_3)$	familiarity weighting	weighted probability		response probability (×1/.4174)	observed probability
cad	.079	0	.0		-	-
ead	.087	0	.0		-	-
eed	.450	.90	.4050		.9704	1.00
end	.019	.50	.0095		.0228	-
oud	.001	1.00	.0010	.4174	.0024	-
vod	.002	.93	.0019		.0045	-

recognition model is shown in Table 1 for one of the words presented in Bouma's recognition study(1973) and its predicted and obtained responses.

When some words in the vocabulary are less well known, they will compete less with better known alternatives on a word recognition trial. Introducing variable word knowledge thus to limit response competition and to counter underestimation of correct response probability by the model.

This is also expected to hold for unfamiliar words, because letter information is usually strong enough to compensate for moderate familiarity. The word vocabulary was thus extended to all 713 Dutch three-letter words, all of which were assigned their familiarity value obtained in the lexical decision study (Chapter 3).

RESULTS OF THE RETEST

The recognition model was tested with the enlarged vocabulary on the same data as in the first two tests with 409 and 541 word vocabularies respectively.[1] Since the letter combination 'ij' was not present in the words serving as stimuli in the lexical decision task, predictions for words containing 'ij' were omitted. There was one of those words in each group of 50 stimuli at the four eccentricities (± 1.75°; ± 2.75°). The same analysis which was carried out for the earlier tests was repeated and the new predictions are compared with those obtained with the vocabulary of 541 words.

Correct word scores

Table 2 shows the average correct word scores and their standard deviations, observed as well as predicted. In the

Table 2

Comparison of observed scores and scores predicted by the recognition model with a 541 word vocabulary and a 713 word vocabulary.

		eccentricity			
		-2.75°	-1.75°	+1.75°	+2.75°
	observed	.535	.786	.904	.729
correct scores	predicted (713)	.482	.719	.849	.626
	predicted (541)	.466	.715	.855	.604
	observed	.216	.215	.140	.207
standard deviations	predicted (713)	.193	.203	.125	.174
	predicted (541)	.185	.210	.125	.189

1) Word recognition data obtained by Bouma (1973)

predictions with the 541 word vocabulary the data on words containing 'ij' were removed. It can be seen that the average amount of underestimation with the 541 word vocabulary is 0.08, and that with the variable knowledge vocabulary it is 0.07. In the predictions of the standard deviations of the correct word scores the underestimation is 0.017 for the smaller vocabulary and 0.021 for the larger one.

Letter scores

The letters which are correct in the observed and predicted responses have been scored for both models again and the summary statistics are shown in Table 3.
It can be seen that with both vocabularies the predicted letter

Table 3

Observed and predicted letter scores in word responses

eccentricity		letters 1	2	3
-2.75°	observed	.850	.691	.705
	predicted (713)	.839	.668	.775
	predicted (541)	.839	.650	.773
-1.75°	observed	.928	.851	.875
	predicted (713)	.945	.854	.831
	predicted (541)	.942	.855	.828
$+1.75^{\circ}$	observed	.949	.944	.978
	predicted (713)	.932	.935	.954
	predicted (541)	.934	.934	.955
$+2.75^{\circ}$	observed	.834	.867	.927
	predicted (713)	.818	.777	.920
	predicted (541)	.817	.768	.919

scores are much closer to the observed ones than the correct scores for words. Nonetheless, there remains a small underestimation of 0.015 with the smaller vocabulary and of 0.012 with the larger one. Another measure of the goodness of fit can be given by the correlations between the 12 (3 letters × 4 eccentric positions) observed and predicted values. The correlation is 0.912 with the smaller vocabulary and 0.918 for the larger one, indicating hardly a difference between the two, though the predictions with the larger vocabulary are a shade better. The underestimation of the letter scores is smaller than that of the correct word scores, because correct letters can also appear in error responses, whose probabilities are somewhat over-estimated.

DISCUSSION

From the comparison of the experimental results and the predictions of the model with two vocabularies it appears that the predictions obtained with the larger, variable knowledge vocabulary are consistently, if slightly, more accurate. This increase in accuracy can not be caused by the 172 additional words in the variable knowledge vocabulary since care was taken that all stimulus words and response words obtained in the experiment formed part of the earlier 541 word vocabulary. The increased accuracy, therefore, must be wholly attributed to the differential weighting of words for inclusion in the set of alternative responses for a presented stimulus word. Though the effect is limited, it must be realized that the larger set of words in the vocabulary also gives rise, on average, to more alternatives for a stimulus word, with the effect that correct response probability decreases. It may be noted that when the vocabulary was extended from 409 to 541 words (+32%) the underestimation increased from 0.03 to 0.08 (Chapter 2). When the vocabulary is extended from 541 words to 713 words, again +32%, the underestimation is reduced to 0.07 when the differential weighting is applied. Still, the underestimation is considerable and calls for an explanation. Three different explanations will be discussed below.

LEXICAL DFCISION VERSUS WORD RECOGNITION

The difference between lexical access in word-nonword decisions and in word recognition was stated in the introduction to be the depth of lexical analysis. If lexical analysis is more extensive for word-nonword decisions, then word familiarity values resulting from such a study should overestimate word accessibility in word recognition and more so for words with a low accessibility. It is unknown, however, to what degree word accessibility is overestimated. It is possible to estimate this from the word recognition data themselves. One way would be to vary the familiarity values in the model exponentially and obtain the best firring exponent. This procedure would basically adapt parameters to recognition data, whereas with the present procedure all parameters are derived independently from other experimental results.

Subject differences in word knowledge

In this connection it is useful to note that the 30 subjects participating in the lexical decision study (Chapter 3) were different from those in the word recognition study (Bouma, 1973; Bouwhuis & Bouma, 1979). Therefore, estimated word knowledge might not be fully representative of the subjects in the recognition experiments.

On each trial there will probably be response alternatives which are better known to the subjects in the lexical decision task than to those in the recognition task and vice versa. On average, the probability of correct responses as such will be equally often overestimated as underestimated. But, by the nonlinear nature of the constant ratio rule the net effect will be an underestimation.

Further, differences in word knowledge between subjects in the recognition task lead directly to variations in word recognition probability which, as discussed below, also lead to underestimation.

Subject differences in recognition

Since the underestimation can be considerably influenced by word knowledge as realized in the vocabularies of the model, shortcomings of the vocabularies seem to be a probable cause of the effect. However it need not be the only one. It has been pointed out earlier (Chapter 2) that differences in letter recognition probabilities between subjects have the effect of underestimating the probabilities of two and even more of three letters. This effect is even strengthened by the limited number of observations for each letter per subject. There were many letters that were presented only twice to all 11 subjects, resulting in 22 observations per letter. Such data preclude an analysis for individual subjects and, moreover, they are prone to large stochastical fluctuations. As a consequence an underestimation is necessarily present in the predictions because of these stochastic and subjective variations. Regarding the number of letters and subjects it is difficult to derive general expressions for the underestimation since the number of unknowns increases rapidly with the number of factors inducing the variations. The only way out is to increase the number of letter recognition trials and to match subjects on recognition performance for letters and for words.

CONCLUSION

The application of the present model with a variable knowledge vocabulary creates somewhat of a precedent in that word frequency data were not used at all and that, instead, word familiarity data were estimated from a reader population. The main advantage of word frequency data is their easy availability, except, of course, that of rarely occurring words. In contrast, familiarity data, as established from lexical decision, may be expected to reflect word accessibility in a more realistic way and by definition can account for subjective knowledge variations. Within the constraints of the model, familiarity data allow the largest possible real word vocabulary. Finally, frequency data do not yield an expectation or probability value directly, whereas familiarity does. The sheer number of possible words, however, seems to preclude an assessment of familiarity for all possible, or a representative part of the words of the language, to which is added the problem of subjective word knowledge differences.

Basically, the operation of variable word knowledge in the letter confusion model is the same as for word frequency in the logogen model (Morton, 1969) and the multicomponent model (Rumelhart & Siple, 1974). In the logogen model less information is needed to trigger the response of a more frequent word; the way in which this is accomplished in the model is treated in the appendix. The effect of word frequency on the information state of the logogen is additive, but since the response probability is an exponential function of the value of the information state, sensory and frequency parameters multiply to form the recognition probability.

The word recognition model of Rumelhart and Siple (1974) incorporates both word familiarity in a crude way and word frequency. First, it is assumed, as a consequence of the experimental paradigm that the subject decides whether the presented string is a word, a syllable or a meaningless string. Within each of these categories a monotonic function of frequency is assumed to combine multiplicatively with the probability of sensory information from a given string. For words this function employs the common word frequency, for syllables the probability of the letter sequence, constrained from left to right is employed. No differential probability is assumed for meaningless strings. In all three models the incorporation of differential word accessibility is, therefore analogous, though accessibility is defined differently and obtained in another way. On the basis of the present analysis it appears useful for recognition and reading studies to obtain a much better insight into human word knowledge and its functional use. This might ever gain more in importance for studying the role of context during reading, where other information constrains the meaning of a specific word. Regarding the difficulty of obtaining sufficiently reliable and representative familiarity data, the finding that positional letter frequency - or spatial redundancy - is a factor in word accessibility is important.

REFERENCES

Boring, E. G. A history of experimental psychology, New York: Appleton Century Crofts, 1950.

Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. Vision Research, 1973, 13, 767-782.

Bouwhuis, D. G. A model for the visual recognition of words of three letters. In: J. Requin (Ed.) Attention and Performance VII, Hillsdale: Erlbaum Ass., 1978.

Bouwhuis, D. G., & Bouma, H. Visual recognition of three-letter words as derived from the recognition of the constituent letters. Perception and Psychophysics, 1979, 25, 12-25.

Bouwhuis, D. G. Word knowledge and letter recognition as determinants of word recognition. In: P. A. Kolers, M. Wrolstad and H. Bouma (Eds.) Processing of visible language I, New York, Plenum Publishing Corporation, 1979.

Bouwhuis, D. G., Schiepers, C. W. J., Schroder, U. O., & Timmers, H. Temporal structure of visual word recognition responses. IPO Annual Progress Report, 1978, 13, 83-93.

Broadbent, D. E. Word-frequency effect and response bias. Psychological Review, 1967, 74, 1-15.

Erdmann, B., & Dodge, R. Psychologische Untersuchungen uber das Lesen, Halle am Saale: Niemeyer, 1898.

Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

Rumelhart, D. E., & Siple, P. Process of recognizing tachistoscopically presented words. Psychological Review, 1974, 81, 99-118.

6 A Review of Related Theories and Experimental Data

This chapter summarizes the basic relations of the letter confusion model with the multicomponent model and the logogen system, which are the only explicit and quantitative models of word recognition. Of these, the logogen system is considered to be equally applicable to auditory word recognition. A further development, notably the cohort theory of spoken word recognition by Marslen-Wilson (1978), shows also interesting correspondences with the present views of constituent letter recognition and lexical access.

The essential visual basis of the letter confusion model permits generalization to a great number of visual recognition experiments. It can provide satisfactory explanations for performance in whole or partial report tasks with the concepts of lateral interference and redundancy.

The effect of word frequency is qualified with the concepts of word familiarity and word knowledge.

Many recognition experiments fall short in describing behaviour representative of normal reading. This seems to be especially true of lexical decision experiments.

The experimental results presented here are thought to yield improved data on word knowledge and lexical access.

Finally, the applicability of our results to reading education is discussed. It may be noted that evaluative research on reading programs is scarce or non-existent. It is argued that there is a particular need for insight into the temporal structure of reading, in addition to assessment of the development of word vocabulary. Furthermore, error analysis, preferably in the framework of a quantative model, may also produce useful information on the reading process in development.

Parts of this chapter have been taken from:
Bouwhuis, D. G., & Bouma, H. Visual recognition of three-letter words as derived from the recognition of the constituent letters. Perception and Psychophysics, 1979, 25, 12-25.

In Chapter 1 reading has been introduced as a frequent and important human activity which seems easy for those who can read but well-nigh impossible for those who cannot.
This state reflects the fact that most processes enabling people to read are not directly observable and are difficult to infer. As a consequence, insight into reading can proceed only by small and seemingly unrevealing steps. The way chosen in this study combines an experimental and a theoretical approach to the issue of single word recognition. Recognizing single words is only a first step in reading, but it is a basic step. The relevance of the theoretical approach is that it may combine a large number of unconnected findings into an integrated whole which may provide the link between recognition of simple features and word perception. The empirical context was made to approximate normal reading situations as closely as is possible with single words.

The basic assumptions for the word recognition model have been reviewed in the introductory chapter. Here it will be shown how the theoretical model developed here relates to other proposals and, on the empirical side, what the relevance is of the findings for reading research and reading education.

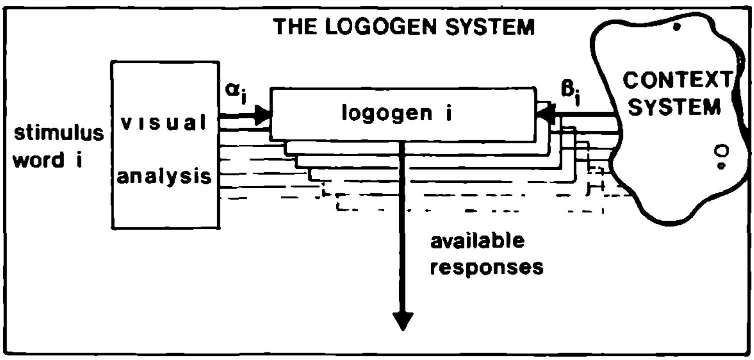


Figure 1. Schematic diagram of the logogen model. The logogen is a device which accepts information from the sensory system concerning the stimulus properties (α_i) and from context producing mechanisms (β_i).

ALTERNATIVE WORD RECOGNITION MODELS

Recently proposed models for visual word recognition include Morton's logogen model (1969) and the multicomponent model put forward by Rumelhart and Siple (1974). The logogen model does not contain a specification of word or letter confusion but the multicomponent model does.
Both the latter and the present letter confusion theory can predict responses for arbitrary words.

The logogen system

The basic unit of Morton's model (1969) is the logogen, (Fig.1) which accepts information concerning a word. During visual processing of stimulus words the logogen is fed information from two sources, the visual system, providing sensory information and the context system, which raises expectations for particular words. The effects of sensory information, α, and context information, β, combine independently, in a multiplicative way, to form the responses strength $\alpha\beta$ for a given word. During reading the response strength is dependent on context, which causes variability in the factor $\alpha\beta$ and consequently in the response strength.
Predictions of responses are made in the framework of choice theory (Luce, 1959). Specifically, the probability of a response r_i is equal to its strength $(\alpha\beta)_i$ divided by the sum of all other response strengths:

$$Q(r_i|s_j) = \frac{(\alpha\beta)_i}{\sum_k (\alpha\beta)_k} \qquad (3)$$

which is equivalent again to the Constant Ratio Rule.
The formula (3) is similar to (3) in Chapter 2, defining the decision rule for the letter confusion model. But the terms $(\alpha\beta)_k$ also show a formal correspondence with the present model.

In experiments concerned with the recognition of single, unconnected words the influence of the context system is minimized. In such a case the stimulus and the visual system are the main sources of variability in the response strengths, allowing the value of β to be set to 1. Now, in (Chapter 2) recognition information from the constituent letters is integrated by way of multiplication too. Its formal counterpart in the logogen model would be $(\alpha_1\alpha_2\alpha_3)_k$, the subscript denoting the three letters, whereas the only difference compared with (2) is that the response strength is only unique up to multiplication by a positive constant. But though the equations are formally identical, the interpretation differs for the two models. The response strength α is originally defined for the whole stimulus word and not for the constituent letters. In the present proposal measurements on the separate letters represent a kind of sensory interpretation of the α factor within the framework of the logogen model.

The original version of the logogen model is rather more detailed with respect to availability of responses. There is generally a parameter V combined with the response strength $(\alpha)_k$, like $(\alpha V)_k$, reflecting the effect of word frequency.

Tests of the logogen model have mainly centered on availability effects and influence of context (Morton, 1969). Disregarding differences in availability it should finally be noted that logogens by definition correspond to real words only. In this way selection of real words from possibly activated strings, as is required in the letter confusion model, is automatically accomplished.

The multicomponent model

The word recognition model proposed by Rumelhart and Siple (1974) encompasses processes responsible for the development of a sensory image of the stimulus word up to the final production of the response word. Here, only the stimulus description and the decision rule will be briefly discussed.

Letters constituting the stimulus words employed by Rumelhart and Siple (1974) consist of straight line segments, which they call 'functional features'. The probability of detection, t_i, of any one feature f_i is assumed to be independent of the presence or absence of other features. For a particular letter consisting of four features, only the features f_1 and f_3 may be detected in a given trial. The probability q of this subset arising is then given by:

$$q = P(f_1,\bar{f}_2,f_3,\bar{f}_4|l) = t_1(1-t_2)t_3(1-t_4) \qquad (4)$$

where l denotes the letter in the word. Letter confusion is introduced here since the two detected features might also belong to other letters. At this stage it is not letters that are mediating agents for the response but parts of letters.

A consequence of independence of detection is that the probability of detecting feature sets for all three letter positions is:

$$P(F(l_1),F(l_2),F(l_3)|s_j) = (q_1q_2q_3)_j, \qquad (5)$$

where $F(l_i)$ denotes the set of extracted features from the letter in position i in the word s_j. These sets of features F might also have been extracted from other letters in other words. The probability of these would generally be different since other letters contain either more or fewer features and in any case different features.
The final response probability is defined as in the former cases:

$$Q(r_i|F(l_1),F(l_2),F(l_3)) = \frac{(q_1q_2q_3)_ib_i}{\sum_k(q_1q_2q_3)_kb_k} \qquad (6)$$

where b_k is the subject's a priori probability that stimulus s_k will be presented. This may be conceived as analogous to the factor β in the logogen model. If this factor is disregarded here for the same reasons as before, (5) and (6) again show a formal correspondence with 2 and 3 of Ch. 1. However, since (6) is defined for parts of letters the predictions of the multicomponent model employ (6) for all different feature sets which may be extracted from the three letters of one stimulus word. The difference compared with the present model of letter confusion to be noted here is the elementary assumption of Rumelhart and Siple (1974) that letter- or feature position does not influence detectability. Though this might have been applicable in the experimental conditions employed by Rumelhart and Siple (1974) it is not valid in general, certainly not for eccentric word presentation.

Frequency effects enter into the picture in a complicated way, since Rumelhart and Siple (1974) predict responses consisting not only of real words, but of syllables and meaningless strings as well. In a simulation of the model Rumelhart and Siple (1974) tried to account for all responses to 726 words and strings of three letters with known frequencies in the printed language. Predictions of the theory, making use of approximations of the parameters, appear to be representative of gross effects.
Frequency effects figure prominently in the data analysis. Letter confusability, which can easily be derived from the multicomponent theory and which is a purely visual effect, produces variations in recognition probability that are much larger than those produced by differences in word frequency.

Stages of processing

Theories of word recognition are often presented in the form of a series of strictly sequential processing stages. At each stage the input is coded into a more abstract form and the corresponding processing operations have a fixed and invariable order. The interpretation is possible that at a given time only one type of processing is applied to the information, while prior information is not functional.
Such a representation may be instructive for explanations of experimental effects as it permits inference as to which processing stages are involved. In Henderson's (1977) words 'it permits us to promenade through the model as if it were a textbook, considering first the phenomena of letter

recognition and finally those of language comprehension'. There are many problems connected with this sequential and hierarchical scheme. Henderson (1977) notes that if access to the lexicon is exclusively phonetic, homophones cannot be distinguished. Further, the traditional explanation of the word superiority effect (Reicher, 1969), stating that words are processed differently from meaningless strings is at variance with such a model, since it would revert the hierarchical order of processing. The views are somewhat reconciled in Johnson's(1977) pattern-unit theory which is strictly sequential, but it delays operations connected with whole units until after iconic processing, and so is strictly hierarchical to the extent that backtracking is severely impeded.

Strict hierarchical processing is made doubtful in particular by the pervasive influence of constituent letters in word recognition and word knowledge, not only in the present experiments, but also in many other studies reviewed in Chapters 1, 2 and 4. Attention may be drawn here to experiments reported by Massaro (1975) and by Estes (1977) in which it was shown that word context does operate given sufficient time, and that its effects continuously increase with processing time. This supports the view that sensory, or letter information remains available. A critique by Kolers (1975) on the hierarchical serial stages model involves the notion that stimuli are not all put through the same program of analysis. Instead, processing may jump from one type of analysis to another type in a nonhierarchical fashion. Winograd (1973) argues that such a system may be operative in language understanding and proposes the name 'heterarchy'. This involves the notion that a number of concurrent processes are working in a coordinated fashion without being under the primary hierarchical control of one of them. The 'distributed memory' model of Hunt (1973) has the same properties for sensory processing. Hunt (1973) cites evidence that e.g. pitch information (high/low) and phoneme information (/ga/-/da/ distinction) access different locations in long-term memory. Yet, also in Hunt's model the processing of stimulus information, as it is controlled by long-term memory information, takes place in serial and discrete sensory buffers.

If letter information remains available at least up to the point of word identification the perceptual information state may be thought to change gradually and not in discrete stages. It is likely that word knowledge information can operate at different points in sensory processing, leading to a response whenever there is sufficient certainty regarding word identity. When there is continuous and simultaneous activity on different levels, say the levels of features, letters and words, the term 'homogeneous' processing seems a better descriptor than either hierarchical or heterarchical[1].

Homogeneous processing may have two aspects. First, information from the stimulus word may be extracted in parallel from the whole string of the constituent letters. If sensory analysis leads to the detection of letters, a content-addressable memory would allow for an efficient access to words matching the letter information available.

As was noted in Chapter 1 letters could well be represented as units in memory, having their own logogens, like words. The operation of letter context - the other letters in the word - could then be analogous to the operation of word context information for word logogens. Letter context would, however, be effective only if some word information is present which could be provided by word context, spatial redundancy, word length or word contour (Estes, 1977).

The second aspect, then, of homogeneous processing involves the nature of the sensory information which is available. Feature information may lead to the recognition of a letter

[1] It must be realized that terms like hierarchical, heterarchical, homogeneous, and also parallel and serial do not uniquely refer to a particular, strictly defined processing model but rather to features of such a model. However, the difficulties in developing a sufficiently detailed and psychologically valid conceptual framework are considerable (Winograd, 1973; Hunt, 1973).

in a given position. In all three models discussed here, the letter confusion model, the logogen model and the multi-component model, words containing that letter in the same position would become more probable as a response. The recognition of letters thus links feature detection directly to word percpetion.

The continuous presence of various types of sensory information has an interesting effect on parallel processing for different words.

In imperfect stimulus conditions several alternative words may fit the available sensory information. In general, the relevance of the sensory information will differ among words and so may the relevance of information of context and word knowledge. Continuous availability of information of all kinds would optimize the selection of appropriate responses, both positively in retaining proper alternatives and negatively in rejecting mismatches.

Imagine, for example, a word for which there is some relevant letter information, but which has a low familiarity. Letter information will then continue to be present in the additional time which is needed to access the lexical entry.

Parallel processing has also been proposed for the *cohort theory* of auditory word recognition by Marslen Wilson (1978) who calls it *distributed processing*, applicable to individual words and morphemes.

Interestingly, the access procedure in the cohort theory is analogous to the content-addressable memory scheme developed in Chapter 4 and is called *pattern directed invocation*. Whereas the visually presented word has a spatial structure which provides simultaneous information from the whole string of letters, the sound signal in auditory presentation has a temporal structure. Marslen Wilson (1978) notes that word knowledge information may be available very soon in processing, indicating that small units of spoken words may already evoke lexical representations Consequently, there is evidence from both visual recognition

and auditory recognition of words in support of a common system of word perception and lexical access.
It is interesting to note that the results pertaining to these conclusions were obtained, like ours, in lexical decision studies and phoneme recognition (phoneme monitoring).

A REVIEW OF RELATED EXPERIMENTAL DATA

Visual interference and letter recognition

The most outstanding difference between the present model and the other recognition models discussed is the incorporation of visual interference. The incorporation is implicit; no theory for the specific effects of interference has been developed here. Its effects are taken to be position specific and are maintained in the model by distinguishing between letters and their positions.

The logogen model does not contain sensory assumptions, but in view of the considerable differences in recognizability resulting from letter position and from word location in the visual field, visual factors might be at least as important as effects of word frequency or expectancy.

The multicomponent model explicitly states independence between feature detectors irrespective of position, though this property has not been tested in the data reported by Rumelhart and Siple (1974). In their experimental conditions, however, the effects of visual interference may have been minimized. The large words, subtending $1.3^\circ \times 2.8^\circ$ visual angle, made up of capitals were presented foveally for durations of the order of milliseconds.
It is a well documented fact that the number of elements, e.g. letters, in a visual display reduces the legibility of each of them. This is especially the case in the parafoveal field and tends to become worse when the elements are closer to each other (Bouma, 1970). Eriksen and Rohrbaugh (1970) found that recognition accuracy decreased progressively when distance between elements was varied from 0.78° to 0.08° visual angle. For the typeface used in the experiments which

are discussed here the average distance is even 0 05° visual angle.
Townsend, Taylor and Brown (1971) found severe reduction of visibility of letters in strings of eight with unlimited viewing time.

The effect is also apparent in the time subjects need to decide on the identity of a designated letter between others. Eriksen and his associates propose one interpretation, holding that the other elements in the display interfere with, or slow down the processing of the target element (Eriksen & Hoffman, 1972; Colegate, Hoffman & Eriksen, 1973; Eriksen & Eriksen, 1974).
Estes (1972) has argued that decreased visual performance which has been interpreted in terms of shifting or focussing of attention can be accounted for in a more coherent way by visual interaction. His interactive channels theory has been formulated quantitatively by Wolford (1975), but not on the level of letter confusions. In this respect Wolford (1975) presents data reported by Hollingworth and Wolford showing that under almost the same conditions as Bouma's (1973), except for a short presentation time of 40 ms, interference starts within 1 degree from the fixation point. Recently Eriksen and Schultz (1977) have shown convincingly that it is mainly the elementary properties of the visual system that affect processing rates for stimuli in the visual field. Response times for single letters increased sharply and linearly with eccentricity where visual acuity decreases. Correspondingly, when the stimulus is degraded, response time are even further increased. On the other hand positional uncertainty does not have any effect. The rate of information extraction is not uniform over the visual field. This rate is then more reduced by the presence of other stimuli, as is the case for words having more letters. This effect is also clear in the whole or partial or single report paradigm for tachistoscopic recognition. This task is in several respects similar to the letter recognition tasks discussed here, at least when the presentation duration does not

exceed 200 ms. Such experiments employ typically eight unconnected letters in a row with the fixation point in the middle of the string. The function relating accuracy of report to stimulus position in the string is W-shaped, but there is a strong interaction with order of report or order of cognitive scanning. This causes the performance enhancement on the left hand side normally encountered (Bryden, 1966; Haber & Standing, 1969; Smith & Ramunas, 1971; Merikle, Coltheart & Lowe, 1971; Schwantes, 1978).
The W-shape found is completely in line with the present findings on interference operative in letter recognition in strings. In the middle of the string near the fixation point foveal acuity is highest and therefore the middle letters can be better discerned and reported. The outward letters are least affected by lateral interference and are consequently more accurately reported than the more inward ones, even though these are closer to the fovea. Such a sensory effect is expected to hold very generally, it is also apparent when two different words or nonwords are presented on each side of the fixation point as in the experiments of Krueger (1976). In his experiments increased response times for letter positions two and three bear witness to reduced sensory information at those positions. Haber and Standing (1969) could stress their argument for less metacontrast by showing that accuracy of report for the end items dropped substantially when the array was prededed and ended by parentheses.
The same type of interference appears in Figs. 1 and 4 of Chapter 2; at greater eccentricities decreasing acuity and increasing interference makes words less recognizable. Initial and final letters are less subject to interference, still it is a curious phenomenon that the letters furthest removed from the fovea are best recognized.
The robustness of these sensory effects, the foveal acuity effect and lateral interference effects (Bouma, 1970; Merikle, Coltheart & Lowe, 1971), even after report delays of up to two seconds (Smith & Ramunas, 1971) seems to leave

little room for the effects of control mechanisms (Schwantes, 1978) except for the apparent left-to-right processing when about eight stimuli have to be recalled.

Retinal locations

In eccentric vision, recognition decreases with eccentricity, but for words the decrease in the right visual fiels is less than in the left. This right field advantage has been shown for English words (Mishkin & Forgays, 1952) and for Dutch words (Bouma, 1973), but not for Hebrew, where, on balance, little left-right difference has been found (Orbach, 1967). It seems, then, that a certain relation exists with the direction of reading. Usually, the right field advantage is taken to reflect the general language specialization of the left cerebral hemisphere, to which the right visual field at first projects. A relation is then assumed with the advantage of the right ear for speech (Kimura, 1961; Bakker, 1970) and with the localization of speech centres mainly in the left cerebral hemisphere (Penfield & Roberts, 1959; Geschwind, 1970).

The present model indicates that the right field advantage for words of three letters is already fully expressed at the level of the constituent letters. Therefore, rather than relating to a general language specialization at levels of encoding more complex than words, these results point to a more efficient coding at a less complex level, perhaps even before the level of letter recognition.

The phenomenon that the right field superiority is higher for words than for the constituent letters is easy to understand in the framework of the letter confusion model. Even if single letter report from meaningless strings is only slightly more accurate in the right field, the combination of all three letters will be much more probable in the right field than in the left field. In the right field, then, erroneous word alternatives will have a lower probability, and there will be fewer of them.

The sum of word probabilities in the right field will there-

fore be only little higher than that in the left field. As a result, when the constant ratio rule is applied the normalizing factor (= 1/(sum of word probabilities)) for the right field will be little smaller than that for the left field.
Though this diminishes to some extent the right field superiority of the string probability, a larger effect than that for single letters can still be observed. Slight effects on parts may therefore have a larger effect on wholes by simple rules of independent combination. Regarding this fact, it is interesting to note that positional letter frequency contributed substantially to word familiarity (Chapter 4). This is another argument for the view that a specialized language system may indeed employ letters in their position for the coding of words.

Completion and redundancy

Middle letters in reported words are much more often correct than when they are reported from meaningless strings. As mentioned before, this is thought to be the completion effect: letters which have not been completely perceived may be inferred from knowledge of the word. From a perceptual point of view word knowledge serves to increase redundancy, which need not be limited to words. One of the basic studies in this field using a recognition paradigm is the experiment by Colegate and Eriksen (1972). They varied several forms of redundancy in letter strings (whereby two letters always appeared together in strings of three) and found a higher proportion of correctly reported letters in redundant strings. The particular type of redundancy only had slight effects. Trying to describe the redundancy effect with a simple rule, Colegate and Eriksen (1972) showed that a perceptual independence model could describe the results satisfactorily. This means in the present context that each letter in the string provides perceptual information which is independent of the information from other letters. This independence condition is the

same as that employed in the present model.
The stimulus conditions in their experiment resemble those in the word recognition experiments reported here and differ only in minor details. The phenomenon that redundancy contributes most under reduced discriminability conditions, according to results cited by Colegate and Eriksen (1972), is clearly borne out by the data of Fig. 4. (Chapter 2)
The representation of word knowledge has been extensively treated in Chapters 3, 4 and 5. It has been assumed throughout that subjects reject nonwords as responses in a word recognition experiment, so that all vocabularies fail to explain the occasional report of nonsense words.
As a rule these responses are pronounceable, suggesting that articulatory readiness could be responsible. It might then be that the perceived letters looked familiar, or sufficiently wordlike to evoke their response. Most of them are predicted by the model to be visually probable. Their rarity, however, impedes a more detailed study.

A few erroneous responses were reported earlier by the subjects in the recognition trials. Sequential word response bias (Morton, 1969) might be responsible for these repetitions when sensory information is apparently limited. Empirically, one could more safely argue in favour of sequential bias if the visual predictability of the response could be assessed more reliably. Thus, the recognition model might be employed to study sequential word bias more accurately.
Finally some words were reported which were not predicted by the model, or predicted to be very improbable. This finding must be expected when the number of observations in the letter recognition study is limited. When only few observations are run the likelihood that a particular letter confusion, which has a low a-priori probability, does not occur at all is rather high. Consequently no word responses would be predicted containing that confused letter, but they might occasionally be reported by a subject.

Frequency effects

Ever since Solomon and Postman (1952) word frequency has been the most extensively studied determinant of word recognition. Apart from its dynamic appeal the universal interest in frequency stems from its basic simplicity as a single variable factor. It is therefore not surprising to find frequency as one of the central issues in the logogen model and the multicomponent model.
Generally word frequency is expressed as number of occurrences in printed text and it is assumed to reflect the number of times a subject encounters that word. More frequent word should then be easier to recognize. However, it is difficult to establish which these frequent words are for any given individual, even when the supposed impact of frequency is valid. In the two word-frequency counts that were available to us, the De la Court count from 1937 (Linschoten, 1963) and that of Uit den Boogaart (1975) differences were of course to be expected on the basis of the forty years separating them, though the kind of material sampled was basically the same. In both counts the less frequent words, of the order of 10^{-6}, counted once out of a total of a million, necessarily have a highly unreliable frequency estimate.
Unreliabilities are also apparent from substantial frequency differences between the same words in printed and spoken text (Uit den Boogaart, 1975), even for frequent words. Taken together these observations seem to indicate that frequency counts inadequately reflect exposure of words to individual subjects, let alone word availability.

An important question is how large the frequency effect is, as compared to ever present visual effects as in reading. Under reasonably favourable perceptual conditions visual effects predominate over those due to frequency, as may be gathered from the data reported by Morton (1969) and by Rumelhart and Siple (1974).
Here an attempt has been made to show how much of word recognition behaviour can be explained without taking

differences in word frequency into account. Even so, the adopted vocabulary represents the basic frequency aspect: the existence of words.

In the lexical decision study (Chapter 3), it was argued that familiarity rather than objective frequency of occurrence reflected the degree to which words are known and the ease with which they are recognized. The only instance in which word frequency had a significant effect was the increase in reaction time, accompanied by a decrease of correct identification for words occurring less than 7 per 10^{-6}. But among these there were many unfamiliar words, which many subjects did not know. Subjects who knew these words needed more time to identify them. Word frequency, or rather infrequency, could have played only a minor role. The absence of a frequency effect for Dutch three-letter words was also noted by Schiepers (in preparation) who presented words varying from 1 to 10 letters in a recognition experiment. For all other wordlengths a moderate frequency effect was obtained both in accuracy and in response time. It may be noted that the Dutch three-letter words constitute a relatively large part of all (26^3 = 17.576) possible three-letter strings; slightly more than 4%. For longer words this fraction decreases rapidly, so that words of more letters have to be known 'better', implying that a frequency effect might be productive.

Word knowledge

One of the most striking observations in the lexical decision study (Chapters 3 and 4) is that subjects made so many errors on words. Strictly speaking, they did not know, 30% of the words presented, which exceeds by far any error rate published in the literature on lexical decision. Stanners, Forbach and Headley (1971) observed an error rate of about 5%; as a rule errors are not even mentioned at all. The high error rate has a curious consequence on the balancing of words and nonwords of the experiment. The ratio of words to nonwords was about 48:52, but since the subjects

missed several words, which were then nonwords for them, this ratio shifted to 34:66 on average, and was even more unbalanced for less well performing subjects. Despite the strict instruction, this could have biased the subject to give nonword responses.
A bias effect of this size, however, is very improbable. It is much more probable that the absence of context and the strict instruction implying that a word meaning had to be found were responsible for the errors. On the other hand subjects in lexical decision tasks are usually given little opportunity to make errors on the small and often carefully chosen sets of words and nonwords. Especially the low number of words may give rise to effects which are specific for that group, a point brought to notice by Clark (1973).
If word frequency is considered to be a crude assessment of familiarity all lexical decision studies corroborate the familiarity effect, except that the effect of word frequency is relatively small, ranging from 70-150 ms. Still, it was found in a review by Whaley (1978) to be the most powerful independent variable, next to measures accounting for richness of meaning, which may also be expected to be related to familiarity. Furthermore the excellent fit of the theoretical familiarity distribution of words (Chapter 3) serves to illustrate the wide variability in word knowledge.

It may be noted here that an analysis of the effect of homographs and of word types (verbs-nouns) was omitted. Dutch three-letter words have frequently more than one meaning, often differing systematically. Essentially, the concept of homography requires a proper definition of what its constituents are: the exact words with different meaning. There is considerable disagreement in the literature as to what a different meaning is, and in fact the Dutch three-letter words presented too many difficulties for a simple analysis of homography.

An analogous situation applied to word type effect. It was expected that verb forms would be more difficult to identify than nouns in isolated presentations. Often three-letter words are both verb and noun, but if separate groups could be formed, frequency or letter composition differed substantially. If it is really true that an isolated verb form is less identifiable than a noun, it is interesting to study its identifiability in running text. Context effects, then could compensate for this lack of identifiability, which would make verbs good candidates for the study of context factors.

The fact that word knowledge is mediated partly by the frequency of the constituent letters in words is an interesting finding which, as it seems, has direct relations with spelling ability. Mason (1978) goes as far as to say that one may know words without knowing the language. This is because letter composition can be learned independently of word meaning. An instructive example in the case of Dutch can be given by means of the rather universal word ski. This Norwegian word of Icelandic origin does not look Dutch at all, but is quite well known. Mason (1975) defines a measure of spatial redundancy which can be calculated for all sequences of the letters of ski. It is defined as the sum of the number of words in which each of the three letters appears in its position, where the words have all the same length. Table 1 shows these numbers and the sums for all possible six strings.

Table 1

Spatial redundancy for letter strings composed of the letters ski. The main body of the table gives the number of words in which the 1st, 2nd and 3rd letters appear and the sum of these numbers

	1	2	3	sum
ski (ski)	23	4	26	53
sik (beard)	23	102	69	194
kis (thin ice)	45	102	70	217
ksi -	45	3	26	74
isk -	9	3	69	81
iks -	9	4	70	83

It so happens that the two strings with the highest scores are both words, though 'kis' is very rare and was known by only two subjects.
If all letters were equally frequent in the three positions the spatial redundancy measure would be
($^1/26$ + $^1/26$ + $^1/26$) x 713 = 82.3. This value seems representative of the last three meaningless strings. On this basis 'ski' would appear to be typically not Dutch, in line with climatic conditions.

Frith (1979) provided evidence that sentences preserving the visual, or letter characteristics, of words were easier to read, for younger readers too, than sentences preserving the phonological characteristics. For writing, Frith (1979) found that good readers could be very bad in spelling, which she attributes to the much larger role of phonological encoding in production.

All these results taken together indicate that for large groups of readers letters are powerful attributes of word recognition and identification.
As the results of Mason (1975) show, it may be that this factor distinguishes good from bad readers. Even such a simple observation seems promising in linking up parts of reading research with reading education.

READING RESEARCH AND EDUCATION

Tests to assess reading performance are an integrative part of reading education, but typically analyze the end stage of an activity which the pupil has acquired. Only in the lowest grade is letter knowledge trained and tested; later this occurs only by means of words and clauses. Quality of reading and recitation - the only directly observable behaviour apart from eye movements - is usually the only measure of reading level and/or reading age.
There are essentially two dependent variables to be observed; the time needed for reading a text of given length, and the number of reading errors. Both variables are likely to be

influenced by ways of articulation, which may be independent of reading level.

Psychometric tests for assessment of reading level employ essentially the same measures but do so under better controlled conditions. Nevertheless, their validity seems to be moderate, as appears from test results on dyslectic and control children.

Readability research

Of course the usefulness of traditional test procedures cannot be denied on this score, since reading education should eventually result in successful reading. But the scores tell little about the reading process itself. Readability research concerns itself rather more with the study of comprehension by which both good and bad readers as well as easy and difficult texts can be revealed. The only widely applied test for comprehension is the cloze test, in which subjects have to complete the words omitted from a printed passage. Applied use of these tests calls for simple predictors of word completion scores which are used as a measure of comprehension. A very complete account of such procedures has been given by van Hauwermeiren (1977). The predictor variables are usually average word length, average sentence length, and level of syntactic complexity. Consequently they reveal little of the especially linguistic processing which leads to a word completion score and to comprehension, though the scores obtained may be very useful in applied settings.

The temporal dimension in reading

From several studies treated in this work it appears that processing time is an important factor in reading. An elementary observation reveals that speech in reading lags behind the eye by some seconds, a point first noted by Buswell (Woodworth, 1938) who called it the eye-voice span.

The intriguing part of reading by eye seems to lie exactly in that period. It has been shown that background knowledge

of readers and context effects can only operate given sufficient time, and that lexical access takes a varying amount of time. For sensory information time variation can be important as well. A word in the eccentric field, right of fixation, will yield only limited sensory information and this will delay processing. Time constants are such that, when the eye fixates the word in the next fixation pause the surfeit of information is processed so fast that recognition of both word presentations may coincide. This suggests that even subtle differences in timing might upset the sequential processing of words, and perhaps even of letters. It has been shown by Bouwhuis et al. (1978) that time constants are surprisingly robust and reveal themselves in widely differing groups of readers and reading tasks. Even so, Bouma and Legein (in preparation) found that it took dyslectic children about 100 ms longer to identify a word than controls, at different levels of accuracy.
This time could be sufficient to lose critical information from visual memory, which would mean that important information for word identification would be unattainable. As repeatedly demonstrated by Bouma and associates (1977a; 1977b) this also leads to decreased performance on letter recognition, expecially in longer words but in meaningless strings as well.

Single word recognition represents basically a static part of the dynamic word recognition process during reading, and as such seems to require additional analyses of temporal structure before the experimental results can be brought to bear on reading education in a successful way.
In any case traditional reading tests should be modified to look much more closely at errors under much better controlled conditions than are feasible in present-day educational settings.

The temporal dimension in reading education

On a much larger time scale the order of acquisition becomes important in what should be learned first. Pre-war

educationalists stressed the importance of relatively restricted but thoroughly known word vocabularies; the work of Thorndike and Lorge (1944) in this area is commonly known. The count of Dutch words by De la Court (1937) also resulted partly from this interest. Vannes (1938) cites professor Ogden for his compilation of 'Basic English' comprising 1000 words, which it was claimed would permit participation in the discussions of any scientific conference. Such a systematic approach ensures the acquisition of a basic vocabulary, but it is not certain whether this would assist in reading, quite apart from technical difficulties for publishing operations.
Nevertheless, it seems relevant to assess the word vocabulary of beginning readers and to pursue its development throughout the years of schooling. This could also reveal possible individual differences which might relate to reading level. In the framework of the letter confusion model it would be interesting to see at what point a letter string which forms a word becomes a meaningful word for the child who starts to read.

Dyslexia

The exact nature of the problems of dyslectic children is not understood. Yet, it might be possible to learn something more by analysing their word recognition scores in the framework of a sufficiently detailed model, which seems valid for normal readers' word recognition. Such a hope may seem wildly optimistic, but it can be argued that this way is more productive than trying to explain word recognition of normal readers with results from dyslectics. One problem is still that dyslectic children have such problems with reading that an otherwise normal recognition paradigm is too tiring and usually yields very low scores.

Further, a detailed recognition model embodying both structural factors relating to accuracy and temporal factors relating to latency, is still lacking, though there are promising possibilities. It seems at any rate useful to analyse reading errors very accurately and to assess the possible causes, such as a shift of attention or eye movements.

The usefulness of error analysis has recently been demonstrated in studies of phonemic dyslexia, an acquired reading disorder due to brain damage by which responses could be attributed to factors described in the logogen model (Marshall & Newcombe, 1977; Patterson, 1978).

Finally it may be possible to alleviate the reading problems of dyslectic children by overtraining them on a limited set of usable words, maybe with a provision for simultaneous sound presentation. Overtraining would result in a threshold lowering, by which less information and global information would be sufficient to evoke correct responses.

Simultaneous sound presentation, though forbidding for teachers by the iterative character might provide useful and even indispensable guidance in acquisition.

REFERENCES

Bakker, D. J. Ear asymmetry with monaural stimulation: relations to lateral dominance and lateral awareness. Neuropsychologia, 1970, 8, 103-117.

Bouma, H. Interaction effects in parafoveal letter recognition. Nature, 1970, 226, 177-178.

Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. Vision Research, 1973, 13, 767-782.

Bouma, H., & Legein, Ch. P. Dyslexia, a specific reading deficit? Manuscript in preparation.

Bouma, H., Legein, Ch. P., & Rens, A. van. Visual recognition by dyslectic children: response latencies for letters and words. IPO Annual Progress Report, 1977b, 12, 87-91.

Bouma, H., & Legein, Ch. P. Foveal and parafoveal recognition of letters and words by dyslectic and by average readers. Neuropsychologia, 1977a, 15, 69-80.

Bouwhuis, D. G., Schiepers, C. W. J., Schröder, U. O., & Timmers, H. The temporal structure of visual word recognition responses. IPO Annual Progress Report, 1978, 83-93.

Broerse, A. C., & Zwaan, E. J. The information of initial letters in the identification of words. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 441-446.

Bryden, M. P. Accuracy and order of report in tachistoscopic recognition. Canadian Journal of Psychology, 1966, 20, 262-272.

Clark, H. H. The language as fixed-effect fallacy: a critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 335-359.

Colegate, R. L., & Eriksen, C. W. Form of redundancy as a determinant of tachistoscopic word recognition. Perception and Psychophysics, 1972, 12, 477-481.

Colegate, R. L., Hoffman, J. E., & Eriksen, C. W. Selective encoding from multi-element visual displays. Perception and Psychophysics, 1973, 14, 217-224.

De la Court, J. F. H. A. De meest voorkomende woorden en woordcombinaties in het Nederlandsch: verslag van een onderzoek in opdracht van het Departement van Onderwijs en Eeredienst, Batavia (Djakarta): Volkslectuur, 1937.

Eriksen, C. W., & Rohrbaugh, J. W. Some factors determining efficiency of selective attention. American Journal of Psychology, 1970, 83, 330-342.

Eriksen, C. W., & Hoffman, J. E. Temporal and spatial characteristics of selective encoding from visual displays. Perception and Psychophysics, 1972, 12, 201-204.

Eriksen, B. A., & Eriksen, C. W. Effects of noise letters upon the identification of a target letter in a non-search task. Perception and Psychophysics, 1974, 16, 143-149.

Eriksen, C. W., & Schultz, D. W. Retinal locus and acuity in visual information processing. Bulletin of the Psychonomic Society, 1977, 9, 81-84.

Estes, W. K. Interactions of signal and background variables in visual processing. Perception and Psychophysics, 1972, 12, 278-286.

Estes, W. K. On the interaction of perception and memory in reading. In: D. Laberge & S. J. Samuels (Eds.). Basic processes in reading: perception and comprehension, Hillsdale: Erlbaum Ass., 1977.

Frith, U. Reading by eye and writing by ear. In: P. A. Kolers, M. E. Wrolstad & H. Bouma (Eds.). Processing of visible language I, New York; Plenum Publishing Corporation, 1979.

Geschwind, N. The organization of language and the brain. Science, 1970, 170, 940-944.

Haber, R. N., & Standing, L. Location of errors with a post-stimulus indicator. Psychonomic Science, 1969, 17, 345-346.

Hauwermeiren, P. van. Het leesbaarheidsonderzoek, Groningen: Tjeenk Willink, 1975.

Henderson, L. Word recognition. In: N. S. Sutherland (Ed.). Tutorial essays in psychology I, Hillsdale: Erlbaum Ass. 1977.

Hunt, E. The memory we must have. In: R. C. Schank and K. M. Colby (Eds.) Computer models of thought and language, San Francisco, W.H. Freeman and Company, 1973.

Johnson, N. F. A. A pattern-unit theory of word recognition. In: D. Laberge and S. J. Samuels (Eds.). Basic Processes in Reading. Hillsdale: Erlbaum Ass., 1977.

Kimura, D. Cerebral dominance and the perception of verbal stimuli. Canadian Journal of Psychology, 1961, 15, 166-171.

Kolers, P. A. Memorial consequences of automatized encoding. Journal of Experimental Psychology: Human Learning and Memory, 1975, 1, 689-701.

Krueger, L. E. Evidence for directional scanning with the order of report factor excluded. Canadian Journal of Psychology, 1976, 30, 9-14.

Linschoten, J. De la Court's frequentietelling van Nederlandse woorden. (Rep. no. 6301) Utrecht: Psychology Laboratory.

Luce, R. D. Individual choice behavior, New York: Wiley, 1959.

Marshall, J. C., & Newcombe, F. Variability and constraint in acquired dyslexia. In: H. Whitaker and H. A. Whitaker (Eds.). Studies in neurolinguistics Vol. 3, New York: Academic Press, 1977.

Marslen-Wilson, W. Sequential decision processes during spoken word recognition. Paper presented at the Psychonomic Society Meetings, San Antonio, November 1978.

Mason, M. Reading ability and letter search time: effects of orthographic structure defined by single-letter positional frequency. Journal of Experimental Psychology, General, 1975, 104, 146-166.

Mason, M. The role of spatial redundancy in grapheme recognition: perception or inference. Journal of Experimental Psychology; Human Perception and Performance, 1978, 4, 662-673

Massaro, D. W. Primary and secondary recognition in reading. In: D. W. Massaro (Ed.). Understanding language, New York: Academic Press, 1975.

Merikle, P. M., Coltheart, M., & Lowe, D. G. On the selective effects of a patterned masking stimulus. Canadian Journal of Psychology, 1971, 25, 264-279.

Mishkin, M., & Forgays, D. G. Word recognition as a function of retinal locus. Journal of Experimental Psychology, 1952, 43, 43-48.

Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

Orbach, J. Retinal locus as a factor in the recognition of visually perceived words. American Journal of Psychology, 1952, 65, 555-562.

Patterson, K. E. Phonemic dyslexia: errors of meaning and the meaning of errors. Quarterly Journal of Experimental Psychology, 1978, 30, 587-601.

Penfield, W., & Roberts, L. *Speech and brain mechanisms*, New Jersey: Princeton University Press, 1959.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 1969, *81*, 275 280.

Rumelhart, D. E., & Siple, P. Process of recognizing tachistoscopically presented words. *Psychological Review*, 1974, *81*, 99-118.

Schiepers, C. W. J. *Response latency and accuracy in visual word recognition*. Ms. 351, Institute for Perception Research, 1979.

Schwantes, F. Stimulus position functions in tachistoscopic identification tasks: scanning, rehearsal and order of report. *Perception and Psychophysics*, 1978, *23*, 219-226.

Smith, M. C., & Ramunas, S. Elimination of visual field effects by use of a single report technique. *Journal of Experimental Psychology*, 1971, *87*, 23-28.

Solomon, R. L., & Postman, L. Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology*, 1952, *43*, 195-201.

Stanners, R. F., Forbach, G. B., & Headley, D. B. Decision and search processes in word-nonword classification. *Journal of Experimental Psychology*, 1971, *90*, 45-50.

Thorndike, E. L., & Lorge, I. *The teacher's wordbook of 30,000 words*. New York: Bureau of Publications, Teacher's College, 1944.

Townsend, J. T., Taylor, S. G., & Brown, D. R. Lateral masking for letters with unlimited viewing time. *Perception and Psychophysics*, 1971, *10*, 375-378.

Uit den Boogaart, P. C. *Woordfrequenties in geschreven en gesproken Nederlands*, Utrecht: Oosthoek/Scheltema en Holkema, 1975.

Vannes, G. De woordenschat van het basis-Nederlandsch. *Tijdschrift voor Levende Talen*, 1938, *4*, 164-175.

Whaley, C. P. Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 143-154.

Winograd, T. A procedural model of language understanding. In:R. C. Schank and K. M. Colby (Eds.) Computer models of thought and language, San Francisco: W. H. Freeman and Company, 1973.

Wolford, G. Perturbation model for letter identification. Psychological review, 1975, 82, 184-199.

Woodworth, R. S. Experimental Psychology. New York: Henry Holt, 1938.

APPENDIX

A Mathematical Exposition of the Logogen Model

In the original paper by Morton (1969) a formal treatment of the logogen system is not given. This section illustrates how response strengths are obtained and how they combine. To this end quotations (in italics) from Morton's paper will be translated into their mathematical equivalent. Page numbers refer to the original.

p.167 *The first step is to assign response strengths to all possible responses in a given situation.*

These response strengths will be denoted as v_i.

p.167 *The probability of any particular response becoming available is then given by the ratio of the response strength for that item divided by the total of the response strengths for all the possible responses.*

$$P_i = \frac{v_i}{v_1 + v_2 + v_3 \ldots}$$

p.167 *We are free to scale the assigned values.*

$$P_i = \frac{v_i}{v_i + \sum v_j} = \frac{v_i / \sum v_j}{v_i / \sum v_j + 1} \quad . \quad (i \neq j)$$

Now put $\beta = v_i / \sum v_j$, then:

$$P_i = \frac{\beta}{\beta + 1} \; .$$

This expression can be considered as a description of the logistic function L, which can also be written as:

$$P_i = \frac{e^c}{e^c + 1}, \quad \text{where } e^c = \beta, \text{ and } c = \ln \beta.$$

This function is shown in Fig. 1c; it can be seen that the response probability P_i rises from 0 to 1 when c increases. The logistic function is a cumulative distribution, arising from a probability distribution ℓ:

$$\frac{dL}{dc} = \frac{e^c}{(e^c + 1)^2}$$, which originates as follows:

p.166 *According to the model there is a continuous exchange of information between the Logogen system and the Context system. This activity affects the values of the counts in the various logogens and it is assumed that samples of the values of the counts would be distributed in a way which approximates the normal distribution.*

Such a distribution is shown in Fig. 1a, which is the logistic distribution with zero mean. The probabilities P_1 can easily be derived from this distribution.

p.166 *On average t items of information are required before the corresponding response will be available.*

The probability that this will happen is indicated by the shaded area in Figs. 1a and 1b. In Fig. 1a there is no information relevant to the stimulus; in Fig. 1b the effect of context is illustrated, shifting the distribution to the right by an amount c. This leads to:

$$P_1 = \int_t^{\infty} \frac{e^{v-c}}{(e^{v-c}+1)^2}\, dv = \frac{1}{e^{t-c}+1} = \frac{e^{c-t}}{e^{c-t}+1} .$$

p.167 *This is equivalent in the Logogen model to assigning a value to the difference between the current level of activation and the threshold for every logogen.*

p.168 *The combined effect of the stimulus together with one of the other factors is calculated by multiplying the response strengths of the two effects for each logogen.*

By generalization this leads to:

$$P_1 = \frac{e^{c+s-t}}{e^{c+s-t}+1}$$, where s represents the sensory effect.

Other effects may also enter.

p.169 *The differences in thresholds of logogens in different frequency classes are indicated by the variables* $V_1 .. V_i .. V_n$.

Denoting V_i by f:

$$P_1 = \frac{e^{c+s+f-t}}{e^{c+s+f-t}+1} .$$

Apparently, all information relevant to the logogen has the effect of shifting the noise distribution in a particular direction. There is no way in the model to distinguish between the sources of information; for example f may be interpreted as a lowering of the threshold t, but also as

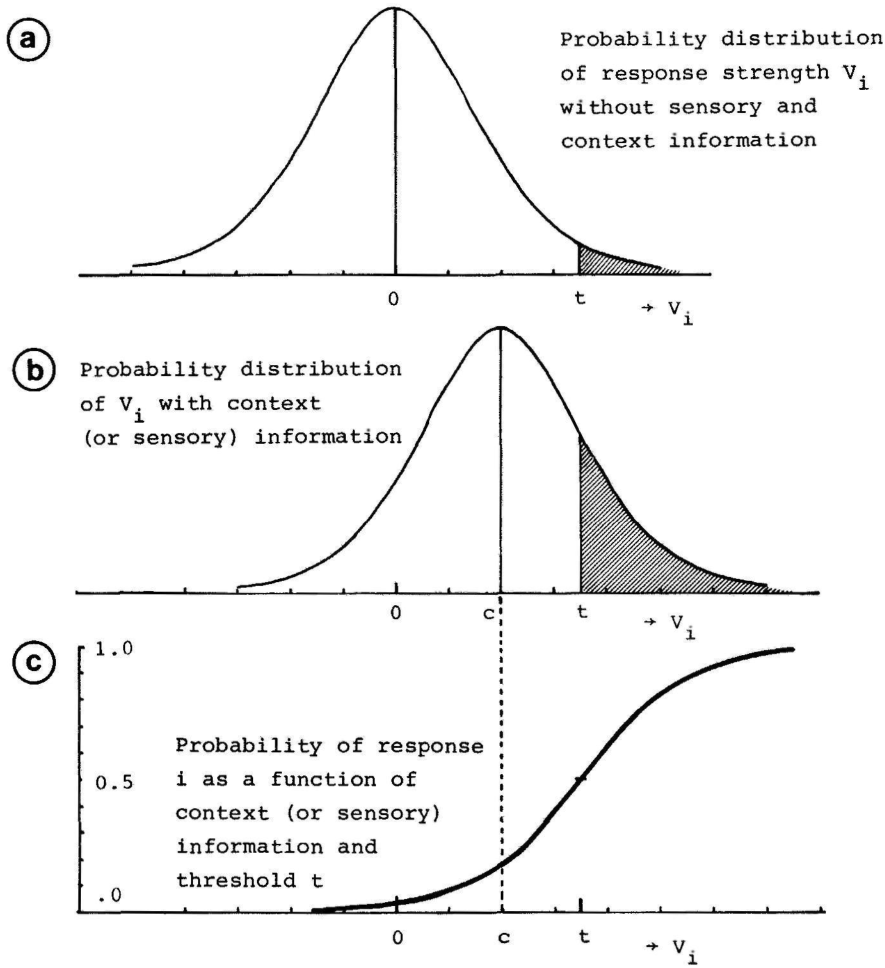


Figure 1. Logistic distributions representing the effects of noise and the effects of information on the values of the counts in the various logogens. All logogens are assumed to have identical distributions. For the distributions of a and b the areas to the right of t are given by the cumulative distribution (bottom) for the values corresponding to the means.

an increase of sensory information added to s.
Finally, it is easy to see how logits directly represent response strength variables in an additive fashion:

$$\begin{aligned} \text{logit } P_i &= \ln \frac{P}{1-P} \\ &= \ln\left\{\frac{e^{c+s-t}}{e^{c+s-t}+1}\right\} \Big/ \left\{\frac{1}{e^{c+s-t}+1}\right\} \\ &= \ln e^{c+s-t} = c + s - t. \end{aligned}$$

REFERENCE

Morton, J. Interaction of information in word recognition. Psychological Review, 1969, 76, 165-178.

SUMMARY

In all the history of mankind reading has probably had the greatest impact on the communication process of man. Successful functioning in present-day society rather stringently requires the ability to read and write.
This does not imply that the world-wide alphabetic writing system is easily mastered. A child can speak, draw or watch television long before he can read. A number of people never learn to read satisfactorily and spelling problems are almost universal.

One of the basic questions tackled in early experimental psychology was how people actually recognize words during reading. As usual, theories split into two opposing camps; analytic recognition holding that letters mediate word perception, and global recognition holding that the whole word functions as a unit. For years these theories also dominated the field of reading education.
Chapter 1 demonstrates that early experimental results on word recognition were wrongly interpreted as unequivocally supporting global recognition during reading.
Subsequent research around 1900 yielded results that were to a varying degree more compatible with analytic perception and these are supported by contemporary research using a wide range of experimental paradigms.

In Chapter 2 a model is proposed for the recognition of words of three letters. This limitation is imposed for practical reasons , and the extension of the model to longer words is dicussed.
The first assumption is that the subject perceives the letters of a presented word probabilistically in their respective serial positions. This gives rise to a number of possible letter strings, some of which will be Dutch words. A second assumption is that only letter strings forming words may be reported by the subject. The probabilities of word responses are generated by a mathematical model on the basis of experimentally obtained data on letter recognition in letter strings resembling words.

Two word vocabularies have been employed for the model. One was derived from the word frequency count by Uit den Boogaart and comprised 409 words. It was extended with additional material from the count by De la Court to give a second vocabulary comprising 541 words. In several respects the word recognition model gives a very adequate description of experimentally obtained results on word recognition. The occurrence of correct responses was, however, somewhat underestimated. It was suspected that this might be due to individual subject's word knowledge being smaller than the vocabularies used in the model.

Chapter 3 presents a lexical decision study designed to derive detailed knowledge of the individual reader's lexicon. For all 713 three-letter words in the largest Dutch dictionary individual subjects' knowledge ranged from 450 to 550 words. Individual words, however, showed a remarkable variation in the degree to which they were known. This word familiarity influenced both decision latencies for correct responses as well as accuracy scores.

Chapter 4 investigates the word properties mediating word familiarity. Familiarity only decreases significantly for very low word frequency. However, the frequency of letters in their position, sometimes called spatial redundancy, is strongly correlated with word familiarity. Digrams are much less effective, so, spelling rules determining or allowing letter sequences are less effective in mediating familiarity than the letter composition of the word.

In Chapter 5 a retest of the word recognition model is described. On the basis of the lexical decision study a new vocabulary was formed in which each of the 713 words was weighted with its familiarity score. Though the increased number of words could have given rise to even more predictions of incorrect responses, familiarity improved the accuracy of the model descriptions slightly compared with the previous tests.

Chapter 6 starts with a comparison of the 'letter confusion' model proposed here with the 'logogen' model of Morton and with the 'multicomponent' model of Rumelhart and Siple. Stochastic independence between stimulus information and word knowledge information are common to all models, in addition, letters are thought to be perceived independently of each other. The 'letter confusion' model, however, incorporates a good deal more of elementary visual factors. It is shown that these are also operative in a large number of different perceptual tasks, stressing their general importance.

The good fit of the 'letter confusion' model, suggesting that letters are basic elements for word perception, supported by the fact that frequent letters play a large part in determining familiarity, lead to the view that letter information and word knowledge may interact directly. 'Bottom up' information is operative for long periods, while 'top down' information can become functional very early in the recognition process. The system proposed for this operation is an associative memory for which the letters of the word function as the address. A similar system with phonemes as basic word constituents has recently been proposed for auditory word recognition.

It is also argued that word familiarity can be used in preference to word frequency for the description of differential word availability.

The present study on word recognition is concerned only with a stationary, though fundamental, part of word recognition during reading. It can be shown that the temporal structure of word recognition is important for ongoing reading. Further research in this area could provide a better link to applied reading research. Approximating the normal reading situation as closely as possible in experimental research might also prove useful for the transfer of empirical results to reading education. Simultaneous print and sound presentation might also be valuable in learning to read.

The appendix treats the mathematical concepts of the 'logogen' model.

SAMENVATTING

Gerekend naar de tijd dat de mens kan spreken en verstaan is lezen een pas betrekkelijk recente mogelijkheid. Toch is de wijze waarop mensen met elkaar communiceren er diepgaand door beinvloed. Het is vrijwel noodzakelijk te kunnen lezen in onze samenleving waar een groot deel van de communicatie verloopt via schrijven, drukken en lezen.
Zulks houdt niet in dat het alfabetische schrift dat bijna overal ter wereld wordt gebruikt, of gebruikt gaat worden, in het geheel geen problemen oplevert. Kinderen leren pas lezen wanneer ze allang kunnen spreken, televisie kijken of tekenen. Een aantal mensen leert nimmer goed genoeg lezen of schrijven om het met vrucht te kunnen gebruiken.

Vragen rond het lezen vormden in de vorige eeuw een der aanzetten tot de experimentele psychologie. Een van de kernvragen, toepasselijk tot in deze tijd, was of woorden herkend worden door middel van de samenstellende letters, ofwel door *het gehele woordbeeld; respectievelijk analytische tegenover* globale herkenning genoemd. Opvattingen ten gunste van een der theorieen hebben ook het leesonderwijs langdurig en wisselend beinvloed. In Hoofdstuk 1 wordt aangetoond dat vroege onderzoeksresultaten, die de globale herkenningstheorie steunden, hoogstwaarschijnlijk onjuist zijn geinterpreteerd. Resultaten van ander onderzoek rond 1900 geven in verschillende mate steun aan de analytische theorie van woordherkenning.
Ook het moderne onderzoek aan leesprocessen levert aanwijzingen dat de samenstellende letters van het woord belangrijke eenheden zijn in woordherkenning.

In Hoofdstuk 2 wordt een model voorgesteld voor de herkenning van woorden van drie letters. Deze beperking berust op practische gronden, de aannames worden geacht geldig te zijn voor andere woordlengtes, zij het dat voor grotere lengtes enige extra aannames nodig zijn. Het model stelt dat de lezer in elke positie van het woord een letter waarneemt, niet noodzakelijk de juiste. Van het grote aantal letterreeksen dat dan waargenomen kan worden, zijn er enkele die bestaande Nederlandse woorden vormen. Aangenomen wordt dat alleen deze in aanmerking komen als antwoord.

Een wiskundig model voorspelt op basis van experimenteel bepaalde waarschijnlijkheden van letterherkenning de kans op het observeren van zowel goede als gespecificeerde foute antwoorden. De twee voor het model gebruikte vocabulaires werden rechtstreeks ontleend aan tellingen van Uit den Boogaart en van de la Court en telden 409 en 541 woorden.In diverse opzichten geven de voorspellingen een uitstekende beschrijving van geobserveerde antwoorden in een eerdere woordherkenningstaak. Correcte antwoorden kwamen echter iets vaker voor dan voorspeld, wat wordt toegeschreven aan een te groot vocabulaire dat teveel foute antwoorden oplevert.

In Hoofdstuk 3 wordt een experiment beschreven dat diende om de woordenschat van de lezer nauwkeuriger te bepalen. Van de 713 drie-letter woorden in de 'van Dale' bleken de proefpersonen er gemiddeld bijna 500 te kennen. Zeer opmerkelijk was echter het grote verschil in bekendheid van de woorden dat tot uiting kwam in de herkenningstijden en in de correcte antwoorden.

In Hoofdstuk 4 wordt beschreven welke eigenschappen van woorden hun bekendheid bepalen. Het effect van woordfrequentie is beperkt tot zeldzame woorden, waarvan de herkenning trager verloopt. Een sterk effect heeft de frequentie waarmee letters voorkomen in hun positie binnen Nederlandse woorden. Voor elke letterpositie geldt dat een daar frequente letter de herkenning sneller en nauwkeuriger maakt. Frequente opeenvolgingen van letters lijken echter niet erg effectief. Veeleer is de schrijfwijze op zich - uit welke letters het woord bestaat - van invloed op het gemak van herkenning.

In Hoofdstuk 5 wordt een nieuw vocabulaire voor het woordherkenningsmodel beschreven. Het bestaat uit alle 713 woorden van drie letters, elk gewogen met hun mate van bekendheid. Hoewel dit vergrote vocabulaire op zichzelf meer foute antwoorden zou voorspellen dan de kleinere vocabulaires, bleek het effect van woordbekendheid de voorspellingen toch enigszins te verbeteren.

In Hoofdstuk 6 wordt het hier voorgestelde 'letter confusion' model vergeleken met het 'logogen' model van

Morton en met het 'multicomponent' model van Rumelhart en Siple. Formele overeenkomsten liggen vooral in de statistische onafhankelijkheid tussen letterherkenningsprocessen enerzijds en letterherkenning en woordkennis anderzijds. Het hier voorgestelde model schenkt echter aanzienlijk meer aandacht aan elementaire visuele factoren, met name aan de toenemende wederzijdse interferentie tussen letters naarmate de afstand van het oogfixatiepunt toeneemt. Dezelfde effecten worden in veel andere visuele taken gevonden en vergroten daarmee de relevantie van visuele factoren bij woordherkenning.

De goede beschrijving die het model levert maakt aannemelijk dat letters in hun positie een belangrijke rol spelen in het herkennen van woorden. Ook het verschijnsel dat letters deels de bekendheid van woorden bepalen suggereert dat letterinformatie rechtstreeks aangrijpt op woordkennis. Een dergelijke werking wordt mogelijk gemaakt wanneer woorden met hun betekenis zijn opgeslagen in een associatief, direct adresseerbaar geheugen. Een dergelijk systeem is onlangs eveneens voorgesteld voor auditieve woordherkenning.

Er wordt verder nader ingegaan op het effect van woordfrequentie, een factor die in de huidige opzet eigenlijk vervangen is door gemeten woordbekendheid.

Hoewel met het bovenstaande alleen een betrekkelijk statisch, maar fundamenteel gedeelte van woordherkenning is beschreven, is bij het lezen het tijdverloop van de woordherkenning van groot belang. Enige argumenten worden gegeven waarom uitbreiding van het onderzoek in die richting zou moeten plaats vinden. Samen met het meer natuurlijk maken van leessituaties in experimenteel onderzoek kan dan wellicht een betere overdracht plaats vinden van onderzoek naar onderwijs. Het effect van gelijktijdige aanbieding van gedrukt woord en geluid zou bij leren lezen bijvoorbeeld vruchtbaar kunnen zijn.

In een appendix worden tenslotte de wiskundige aspecten van het 'logogen' model van Morton nader uitgewerkt en toegelicht.

CURRICULUM VITAE

Dominicus Gabriel Bouwhuis werd 13 september 1942 in Leeuwarderadeel geboren. Na de lagere school bezocht hij vanaf 1954 de Julianaschool voor U.L.O. te Leeuwarden en vanaf 1956 het Stedelijk Gymnasium aldaar, waar hij in 1962 het einddiploma Gymnasium α verkreeg. Aan de Katholieke Universiteit te Nijmegen studeerde hij vervolgens psychologie, waarvoor hij in 1965 zijn kandidaatsexamen behaalde. Daarna volgde specialisatie in psychologische functieleer, mathematische psychologie en wiskunde. Hij behaalde het doctoraal examen in 1968 met als afstudeeronderwerpen beslissingsmodellen in kort geheugen onderzoek en de Simplex schaaltheorie van Guttman. Tussentijds vervulde hij een stageperiode op het Concern Stafbureau van de N.V. Philips over manpower planning. Op 1 augustus 1968 trad hij in dienst op het Instituut voor Perceptie Onderzoek te Eindhoven, gedetacheerd door het Natuurkundig Laboratorium der N.V.Philips. Tot 1977 was hij werkzaam in de groep Perceptieve en motorische vaardigheden en werd daarna leider van de nieuwe onderzoeksgroep Cognitie en communicatie

STELLINGEN

1. De vaststelling van Geyer (1977) dat foveale letterherkenning kwalitatief verschilt van parafoveale berust op nog andere experimentele artefacten dan door Molloy (1978) worden genoemd en op de toepassing van een onjuist model.

   Geyer, L. H. Recognition and confusion of the lowercase alphabet. Perception and Psychophysics, 1977, 22, 487-490.
   Molloy, D. G. Comments on recognition and confusion of the lowercase alphabet. Perception and Psychophysics, 1978, 24, 190-191.

2. Zowel Townsend (1971) als Geyer (1977) gebruiken procedures die tot onjuiste parameterschattingen leiden voor het alles-of-niets herkenningsmodel.

   Townsend, J. T. Theoretical analysis of an alphabetic confusion matrix. Perception and Psychophysics, 1971, 9, 40-50.
   Geyer, L. H. Op. cit.

3. De effecten van klinkerduren op de waarneming van prosodische grenzen in een betekenisvolle spraakuiting zijn onafhankelijk van elkaar.

   Bouwhuis, D. G., & de Rooij, J. J. Vowel length and the perception of prosodic boundaries. IPO Annual Progress Report, 1977, 12, 63-68.

4. Als proefpersonen een gegeven aantal spraakattitudes vrij mogen benoemen, gaat geen selectieve informatie verloren wanneer andere proefpersonen deze beschrijvingen weer herleiden tot de door de sprekers bedoelde attitudes, behalve met betrekking tot de beschrijving 'neutraal'.

   Bouwhuis, D. G. The recognition of attitudes in speech. IPO Annual Progress Report, 1974, 9, 82-86.

5. Als regel wordt bij de toepassing van de binomiaaltoets in experimentele omstandigheden de overschrijdingskans op een geobserveerd verschil overschat vanwege variabiliteit in de kans p. In sommige situaties (van Katwijk, 1974) is het mogelijk die variabiliteit te schatten en daarmee de toets efficienter te maken.

   Bouwhuis, D. G. Het klemtoonoordeel, persoonlijk of eenstemmig? IPO rapport nr. 253, 1973.
   van Katwijk, Accentuation in Dutch. Proefschrift, Rijksuniversiteit Utrecht, 1974.

6. De door Shepard (1972) om practische redenen voorgestelde formule voor de gelijkenis S_{ij} tussen twee stimuli i en j, te ontlenen aan een verwarringsmatrix P:

$$S_{ij} = \frac{p_{ij} + p_{ji}}{p_{ii} + p_{jj}}$$

levert in feite een zuiverder schatter voor de gelijkenismaat η van het keuzemodel van Luce (1959) dan de voorgestelde

$$\hat{\eta}_{ij} = \left\{\frac{p_{ij}}{p_{ii}} \cdot \frac{p_{ji}}{p_{jj}}\right\}^{\frac{1}{2}}.$$

Dit geldt echter alleen als de bijbehorende biaswaarden β binnen de meetonnauwkeurigheid aan elkaar gelijk zijn, hetgeen regelmatig voorkomt.

Shepard, R. N. Psychological representations of speech sounds. In: E. E. David and P. B. Denes (Eds.) Human communication, New York: McGraw-Hill, 1972.
Luce, R. D. Detection and recognition. In: R. D. Luce, R. Bush en E. Galanter (Eds.) Handbook of mathematical psychology I, New York, Wiley, 1963.

7. De formule die Ohala en Lyberg (1976) geven voor de door meetfouten ontstane correlatie tussen de duren van twee opeenvolgende spraaksegmenten:

$$\frac{-VE}{\{(VX + VE)(VY + VE)\}^{\frac{1}{2}}};$$

waarin VX en VY de varianties zijn van de segmentduren en VE de meetfoutvariantie voorstelt; is onjuist voor ingebedde segmenten en moet luiden:

$$\frac{-VE}{\{(VX + 2VE)(VY + 2VE)\}^{\frac{1}{2}}}.$$

Daar zij bij de afleiding van de correlaties evenmin rekening houden met suprasegmentale tijdsstrukturen is hun kritiek op Wright's onderzoek (1974) inhoudsloos.

Ohala, J. J., & Lyberg, B. Comments on "Temporal interactions within a phrase and a sentence context." Journal of the Acoustical Society of America, 1976, 59, 990-992.
Wright, T. W. Temporal interactions within a phrase and a sentence context. Journal of the Acoustical Society of America, 1974, 56, 1258-1265.

(Bij D. G. Bouwhuis, Visual recognition of words)

Nijmegen, Aula Universiteit, Wilhelminasingel 13
6 juni, 1979, 16.00 uur.